

# FX Trading and the Exchange Rate Disconnect Puzzle

Martin D. D. Evans\*

Department of Economics, Georgetown University.

First Draft: Comments Welcome

5th November 2018

## Abstract

This paper examines how trading in the FX market carries the information that drives movements in currency prices over minutes, days and weeks; and how those movements are connected to interest rates. The paper first presents a model of FX trading in a Limit Order Book (LOB) that identifies how information from outside the market is reflected in FX prices and trading patterns. I then empirically examine this transmission process with the aid of a structural VAR estimated on 13 years of LOB trading data for the EURUSD, the world's most heavily traded currency pair. The VAR estimates reveal several new findings: First, they show that shocks from outside the LOB affect FX prices through both a liquidity and information channel; and that the importance of these channels varies according to the source of the shock. Liquidity effects on FX prices are temporary, lasting between two and ten minutes, while information effects of shocks on prices are permanent. Second, the contemporaneous correlation between price changes and order flows varies across shocks. Some shocks produce a positive correlation (as in standard trading models), while others produce a negative correlation. Third, the model estimates imply that intraday variations in FX prices are overwhelmingly driven by one type of shock, it accounts for 87% of hour-by-hour changes in the FX prices.

The second part of the paper examines the connection between the shocks in the trading model and the macroeconomy. For this purpose, I use the VAR estimates to decompose intraday FX price changes and order flows into separate components driven by different shocks. I then aggregate these components into daily and weekly series. I find that one component of daily order flow is strongly correlated with changes in the long-term interest differentials between US and EUR rates. This suggests that the intraday shocks driving this order flow component carry news about future short-term interest rates which is embedded into FX prices. I find that intraday shocks carrying interest-rate information account for on average 56% of the variance in the daily EURUSD depreciation rate between 2003 and 2015, but their variance contributions before 2007 and after 2011 are over 80%. These findings indicate that the EURUSD depreciation rate is relatively well-connected to macro fundamentals via a particular component of order flow. Finally, I show that flows embedding liquidity risk have forecasting power for daily and weekly EURUSD depreciation rates.

**Keywords:** FX Trading, Microstructure, Order Flow, Exchange-Rate Determination

**JEL Codes:** F31, F32, F34

\*Email: [evansmdd@gmail.com](mailto:evansmdd@gmail.com). Mailing Address: Department of Economics, Georgetown University, Washington D.C. 20057.

## Introduction

This paper examines how trading flows in the foreign exchange (FX) market carry the information that drives movements in spot exchange rates over minutes, days and weeks; and how those movements are connected to interest rates. Short-term movements in exchange rates are generally viewed as having little connection to macroeconomic variables, an observation Obstfeld and Rogoff (2000) call the Exchange-Rate Disconnect Puzzle. I argue here that short-term movements in the EURUSD rate (the world’s most heavily traded currency pair) are in fact quite well “connected” to interest differentials, but the connection is made via trading flows. More specifically, I show that a particular flow, identified from a trading model estimated in high-frequency intraday data, accounts for 56% of the daily variation in the EURUSD rate over the 13 years between 2003 and 2015, and that the flow is strongly linked to changes in the differential between one-year US and Euro interest rates. The connection appears to be even stronger away from the 2008 financial crisis. Excluding the 2007-11 period, I find that at least 80% of the daily variations in the EURUSD rate are attributable to flows linked to the interest differential. The trading model also identifies another flow that captures changes in the risks of providing liquidity to the FX market. These risks are quite different from those identified in standard asset-pricing models of the exchange-rate risk premia.<sup>1</sup> I show that flows embedding liquidity risk have some forecasting power for daily and weekly changes in the EURUSD rate.

I begin my analysis with a microstructure model of modern FX trading on an electronic Limit Order Book (LOB). The model identifies different transmission channels through which shocks from outside the trading venue affect trading flows and FX prices. Next, I estimate a structural VAR that quantifies the importance of these different channels using 13 years of EURUSD trading data sampled every 30 seconds. The VAR identifies the components of intraday trading flows and FX price changes that are driven by different shocks, which I aggregate into daily and weekly time series. Finally, I investigate how the daily trading flows identified by the VAR are linked to interest rate differentials, how they connect daily depreciation rates with changes in differentials, and how they contribute to the predictability of future daily and weekly depreciation rates.

This analysis builds on several areas of prior research, including: microstructure models of FX trading, empirical exchange-rate models using trading flows, and the large macro-based literature on exchange rate modeling. Earlier FX microstructure models, such as Lyons (1997), Evans and Lyons (2002), Osler et al. (2011) and Evans (2011), consider trading in a two-tier market. The FX dealers working at major banks trade with each other in the inner tier, while trades between dealers and end-users take place in the outer tier. These models used highly stylized trading protocols to focus on how information embedded in individual trades between dealers and end-users was transmitted across the market via intra-dealer trades. The institutional structure of FX trading has evolved considerably since the original development of these models (see, e.g., King et al., 2011). FX trading still takes place in a two-tier market, but trading in the inner tier, often referred to as the wholesale market, now takes place between dealers and other sophisticated trading entities such as hedge funds, mutual funds, and HFTs. There has also been a proliferation in electronic trading

---

<sup>1</sup>See Engel (2014) for a recent survey.

venues across both tiers. Electronic LOBs run by EBS and Reuters have been the primary wholesale trading venues for spot FX trading since the late 1990s. Trading in the major currency pairs, like the EURUSD, GBPUSD, and USDJPY, are concentrated on one or other of these LOBs.<sup>2</sup>

In light of these institutional changes, I begin my analysis by studying trading flows and price dynamics in a new microstructure model of FX trading on a LOB.<sup>3</sup> The model contains two key features. First, it recognizes that traders have limited information about outstanding limit orders when they make their own trading decisions. Second, it confers a temporary informational advantage on traders placing market orders over those with outstanding limit orders. With these features, the model identifies two distinct transmission channels: a liquidity and information channel. Shocks have temporary effects on FX prices via the liquidity channel, and permanent effects via the information channel. I use the model to examine how different types of shocks to the LOB produce different liquidity and information effects, which in turn generate distinct correlations between changes in FX prices and order flows.<sup>4</sup> These correlations can be positive, zero, or negative, depending on the type of shock.

My empirical analysis is based on estimates of a structural VAR that uses high-frequency (30-second) trading data covering a much longer time span than earlier empirical research on FX trading. The data comprise information on EURUSD spot trades on the LOB run by EBS between midnight on January 6, 2003 and December 31, 2015. Trading data from the EBS system has been previously used by Hau et al. (2002), Killeen et al. (2006), Berger et al. (2008a) Chinn and Moore (2009), Breedon and Vitale (2010), and Ito and Yamada (2015).<sup>5</sup> One distinguishing feature of my empirical analysis is that it uses high-frequency trading data over a time span that covers a variety of macroeconomic conditions, including the 2008 financial crisis. Hasbrouck (1991) pioneered the use of VARs in the empirical analysis of trading data. My use of a VAR to examine FX trading follows earlier work by Payne (2003), Daniélsson and Love (2006), Menkhoff and Schmeling (2008), and others. Here the VAR includes six variables and imposes a non-recursive structure that allows multiple shocks to affect FX prices and trading flows simultaneously. This feature allows the model to accommodate the different co-movements in FX prices and order flows induced by different shocks, consistent with the predictions of the microstructure trading model.

The VAR estimates reveal that two innovations account for almost all the variations in FX prices and order flow in the EURUSD market. One innovation produces impulse responses similar to the effects of exogenous

---

<sup>2</sup>At the time of their study, Breedon and Vitale (2010) estimated that 88% of EURUSD spot trades took place on the EBS venue. In the last few years trading on electronic platforms run by individual banks (e.g. Deutsche Bank's Autobahn and Barclays' BARX platforms) has expanded significantly at the expense of EBS and Reuters. Nevertheless, these LOBs remain the single most important venues for spot trading in the market.

<sup>3</sup>Trading in many other financial instruments such as equities, bonds, and derivatives takes place on electronic LOBs, so there is a substantial body of research analyzing these venues within the broad field of microstructure (see, Parlour and Seppi, 2008 for a survey). Because LOBs allow for the submission of both market and limit orders, finding how the optimal trading strategies of individual traders combine to produce equilibrium dynamics for prices and trade flows is an extremely complex undertaking (see, e.g., Foucault et al., 2013). To avoid this complexity, my model focuses on the aggregate implications of individual traders' decisions for prices and trade flows, rather than the individual decisions.

<sup>4</sup>Order flow is defined as the difference in value between flows of market orders to buy FX and market orders to sell FX.

<sup>5</sup>These studies use several years of trading data at the daily or monthly frequency, or they examine a few years of intraday data. For example, Hau et al. (2002) and Killeen et al. (2006) studied EBS order flow data at the daily frequency from 1998 and 1999. Breedon and Vitale (2010) studied 6 months of intraday data from 2000 and 2001, while Berger et al. (2008a) use six intraday data from EBS on EURUSD and USDJPY trading between January 1999 and December 2004. Ito and Yamada (2015) also use EBS data to study trading around the Tokyo and WMR fixes from 1999 to 2013. Yearly spans of Reuters LOB data have been used to examine the effects of data releases on order flows by Love and Payne (2008) and Rime et al. (2010).

shocks to market orders in the microstructure model. These innovations produce a positive correlation between FX price changes and order flows and account for approximately 87% of the variance in hour-by-hour EURUSD depreciation rates. The other innovation produces impulse responses similar to the effects of shocks to the supply of limit orders. These innovations produce a negative correlation between unexpected FX price changes and order flow and account for approximately 13% of the hour-by-hour depreciation variance. The estimates also reveal that while the liquidity channel is operable, its quantitative impact is small. Shocks primarily affect FX prices and trading flows via the information channel.

These empirical findings are quite novel. Following Evans and Lyons (2002), numerous studies have found a strong positive correlation between order flows and FX price changes measured over periods ranging from a few minutes to a few weeks.<sup>6</sup> My estimates suggest that this unconditional correlation masks the complex effects of different shocks. The results also contribute to the debate about the relative strengths of liquidity and information effects. Papers by Froot and Ramadorai (2005), Berger et al. (2008a) and Breedon and Vitale (2010) argue that liquidity effects are dominant, whereas Evans and Lyons (2002, 2006, 2013), Marsh and O'Rourke (2005), Killeen et al. (2006), Bjønnes et al. (2007) and Love and Payne (2008) stress the importance of information effects.

In the last part of the paper, I use the VAR to identify the different components driving daily and weekly movements in the EURUSD rate and examine their links to interest rates and FX risk. Earlier research by Evans (2010) and Evans and Lyons (2013) found that order flows from end-users trading with Citibank contain incremental (non-public) information about (unreported) current and future macro variables. My results support and extend these findings. In particular, I show that daily order flows driven by intraday market order shocks are strongly correlated with daily changes in the difference between US and EUR one-year nominal interest rates. This is consistent with the idea that these intraday shocks carry new information about future short-term interest rates. Similarly, Rime et al. (2010) find that order flows from the Reuters LOB are linked to the unexpected macroeconomic announcements in 2004. Furthermore, I estimate that on average these intraday shocks (carrying interest rate news) account for approximately 56% of the variance in the daily EURUSD depreciation rates between 2003 and 2015. While this is an unusually high degree of explanatory power for an exchange-rate model, the connection between depreciation rates and interest rates appears to be even stronger away from the 2008 financial crisis. I estimate that the variance contribution of order flow shocks is over 80% in periods before 2007 and after 2011.<sup>7</sup>

If one component of order flow carries information about interest rates, what is the role of the other order flow component? The answer, I argue, is that it reflects changes in the risk of providing liquidity to the market. Shifts in the perceived distribution of future order flows and FX prices alter the risks associated

---

<sup>6</sup>Positive correlations between order flow and returns also appear other asset classes; see, e.g. Hasbrouck (1991) for stock markets and Brandt and Kavajecz (2004) for U.S. bonds. For surveys of the empirical FX literature on order flows, see Osler (2009) and Evans and Rime (2012).

<sup>7</sup>By comparison, exchange-rate models that incorporate macro fundamentals via Taylor-rules for interest rates (see, e.g., Engel and West, 2006, Engel et al., 2008 and Mark, 2009) account for less than 10% of month/quarterly depreciation rates (see Evans, 2011 and Engel, 2014 for discussions). Regressions of daily depreciation rates on order flows estimated over shorter data spans often produce  $R^2$  statistics around 0.5, but these models are silent on what drives the order flows. A partial list of studies using customer flows includes; Marsh and O'Rourke (2005) for major currencies, King et al. (2010), Bjønnes et al. (2005), and Rime (2001) for small open economies, and Gyntelberg et al. (2009), Onur (2008) and Wu (2007) for emerging markets.

with the submission of limit orders. Such shifts should appear as shocks to the supply of limit orders, which drive the second component of daily order flow identified by the VAR. They should also show up in future FX price changes. Consistent with this view, I find that the order flow component embedding shifts in liquidity risk has statistically significant (within-sample) forecasting power for daily and weekly changes in the EURUSD rate that are unrelated to future changes interest rates. This finding adds to earlier results on the forecasting power of order flow in Evans and Lyons (2005, 2013), Bjønnes et al. (2005) and Killeen et al. (2006).

The remainder of the paper is structured as follows: Section 1 presents the microstructure trading model. In Section 2, I describe my intraday empirical analysis using the structural VAR. Section 3 contains the macro empirical analysis connecting the EURUSD rate to interest rates and liquidity risk. Section 4 concludes.

## 1 The Trading Model

I use a simple model of trading on a LOB to examine how different shocks are transmitted to FX prices and trading flows. The model incorporates into a LOB setting aspects of inventory control and asymmetric information found in earlier microstructure models of dealer markets (for overviews, see Hasbrouck, 1996, O’Hara, 1997 and Foucault et al., 2013). One central feature of the model is that traders have limited information about the state of the LOB: they only observe the best prices and the flows of market orders rather than the complete structure of limit orders. There are similar limits on the information available to actual FX traders on LOBs. The model identifies two distinct transmission channels: a liquidity and information channel. My analysis examines how different types of shocks to the LOB produce different liquidity and information effects, which in turn generate distinct co-movements in FX prices and trading flows.

### 1.1 Set Up

Let  $p_t^{\text{off}}$  and  $p_t^{\text{bid}}$  denote the logs of the best offer and bid prices for FX on the LOB at the start of period  $t$ . These prices are identified as the lowest offer price among the outstanding limit sell orders for FX and highest bid price among the outstanding limit buy orders for FX. I assume that these prices depend on the current estimated fundamental value of FX,  $\mu_t$ , and the estimated depth (i.e. total value) of limit orders on either side of the book. Let  $d_t^{\text{sell}}$  and  $d_t^{\text{buy}}$  respectively denote the depth of the sell and buy limit orders at the start of period  $t$ . The best limit prices are determined as

$$p_t^{\text{off}} = \mu_t + \frac{1}{2}\delta - \beta E_t d_t^{\text{sell}} \quad \text{and} \quad p_t^{\text{bid}} = \mu_t - \frac{1}{2}\delta + \beta E_t d_t^{\text{buy}}, \quad (1)$$

where  $\delta$  and  $\beta$  are positive constants. Expectations conditioned on information at that start of period  $t$  are denoted by  $E_t$ .

The value of  $\delta$  determines the size of the spread between the bid and offer prices when the estimated depth of the book is zero, i.e.,  $p_t^{\text{off}} - p_t^{\text{bid}} = \delta$  when  $E_t(d_t^{\text{sell}} + d_t^{\text{buy}}) = 0$ . This parameter captures the adverse

selection cost of submitting a limit order. The  $\beta$  coefficient determines the sensitivity of the best limit prices to the estimated depth. Since market orders are matched against the limit order with the best price, traders submitting a limit order balance the benefit from the order being matched at a higher offer (lower bid) against the cost of their order not being matched. Ceteris paribus, an increase in the depth reduces the probability of a match, so when the estimated buy (sell) depth increases, traders submit buy (sell) orders with higher (lower) prices. The  $\beta$  coefficient determines the degree to which the best limit prices move as changes in the estimated depth alter traders' limit order submissions. In particular, an increase in the estimated depth of sell orders lowers the price of the lowest offer, while an increase in the estimated depth of buy orders raises the price of the highest bid. Note that prices reflect estimated rather than actual depth because traders cannot observe all outstanding limit orders.

The flows of market orders to buy FX and sell FX during period  $t$  depend on the difference between the expected trade price and the estimated fundamental value of FX. Because individual market orders are sequentially matched with the best outstanding limit orders, the actual trade price for an individual market buy (sell) order will differ from  $p_t^{\text{off}}$  ( $p_t^{\text{bid}}$ ) when the volume of preceding orders exceeds the depth of the limit orders at  $p_t^{\text{off}}$  ( $p_t^{\text{bid}}$ ). I assume that the individual market orders submitted in period  $t$  are randomly sequenced so that the (ex ante) expected trade price for an individual order is a weighted average of the current best price and the expected best price next period. Specifically, the expected trade price for market buy and sell orders are respectively  $(1 - \omega)p_t^{\text{off}} + \omega E_t^* p_{t+1}^{\text{off}}$  and  $(1 - \omega)p_t^{\text{bid}} + \omega E_t^* p_{t+1}^{\text{bid}}$  with  $1 > \omega > 0$ , where  $E_t^*$  denotes the expectations of traders placing market orders. The flows of market orders to buy FX and sell FX during period  $t$  are determined by:

$$\begin{aligned} m_{t+1}^{\text{buy}} &= -\alpha \left[ (1 - \omega)p_t^{\text{off}} + \omega E_t^* p_{t+1}^{\text{off}} - \frac{1}{2}\delta - \mu_t \right] + u_{t+1}^{\text{mbuy}} \quad \text{and} \\ m_{t+1}^{\text{sell}} &= \alpha \left[ (1 - \omega)p_t^{\text{bid}} + \omega E_t^* p_{t+1}^{\text{bid}} + \frac{1}{2}\delta - \mu_t \right] + u_{t+1}^{\text{msell}}, \end{aligned} \quad (2)$$

where  $\alpha > 0$ . Here  $u_{t+1}^{\text{mbuy}}$  and  $u_{t+1}^{\text{msell}}$  denote exogenous shocks to the flows of market buy and sell orders, respectively. They comprise a common and directional component:  $u_{t+1}^{\text{mbuy}} = \frac{1}{2}u_{t+1}^{\text{vol}} + \frac{1}{2}u_{t+1}^{\text{flw}}$  and  $u_{t+1}^{\text{msell}} = \frac{1}{2}u_{t+1}^{\text{vol}} - \frac{1}{2}u_{t+1}^{\text{flw}}$ . The  $u_{t+1}^{\text{vol}}$  shock contributes equally to the flows of market buy and sell orders, while the  $u_{t+1}^{\text{flw}}$  shock changes the balance between the flows. The  $u_{t+1}^{\text{vol}}$  and  $u_{t+1}^{\text{flw}}$  shocks will be interpreted as volume and order flow shocks below. Note that market orders during period  $t$  appear with a  $t + 1$  subscript to make clear that they embody participants' decisions to trade after period- $t$  limit prices are determined.

The depth on each side of the LOB depends upon the flows of market orders and the submission of new limit orders. The depth of outstanding limit orders to buy and sell FX at the start of period  $t + 1$  are given by

$$d_{t+1}^{\text{buy}} = d_t^{\text{buy}} - m_{t+1}^{\text{sell}} + u_{t+1}^{\text{lbuy}} \quad \text{and} \quad d_{t+1}^{\text{sell}} = d_t^{\text{sell}} - m_{t+1}^{\text{buy}} + u_{t+1}^{\text{lsell}}, \quad (3)$$

where  $u_{t+1}^{\text{lbuy}}$  and  $u_{t+1}^{\text{lsell}}$  respectively denote the exogenous shocks to the supplies of limit orders to buy and sell FX during period  $t$ . Again I write the limit supply shocks as the sum of common and directional components:  $u_{t+1}^{\text{lbuy}} = \frac{1}{2}u_{t+1}^{\text{d}} + \frac{1}{2}u_{t+1}^{\text{bal}}$  and  $u_{t+1}^{\text{lsell}} = \frac{1}{2}u_{t+1}^{\text{d}} - \frac{1}{2}u_{t+1}^{\text{bal}}$ . The  $u_{t+1}^{\text{d}}$  component equally affects the supply of limit

orders on both sides of the LOB, while the  $u_{t+1}^{bal}$  component changes the balance between the supplies to each side of the LOB.

To facilitate the analysis below it proves useful to rewrite equations (1) - (3) in terms of some new variables. The new variables are the mid-point FX price,  $p_t = \frac{1}{2}p_t^{off} + \frac{1}{2}p_t^{bid}$ ; the spread between the best bid and offer prices,  $spd_t = p_t^{off} - p_t^{bid}$ ; total depth,  $d_t = d_t^{buy} + d_t^{sell}$ ; the depth balance,  $bal_t = d_t^{buy} - d_t^{sell}$ ; trading volume,  $vol_t = mt_t^{buy} + mt_t^{sell}$ ; and order flow  $flw_t = mt_t^{buy} - mt_t^{sell}$ . Rewriting (1) - (3) in terms of these new variables gives:

$$flw_{t+1} = -2\alpha [(1 - \omega)p_t + \omega E_t^* p_{t+1} - \mu_t] + u_{t+1}^{flw}, \quad (4)$$

$$bal_{t+1} = bal_t + flw_{t+1} + u_{t+1}^{bal}, \quad \text{and} \quad (5)$$

$$p_t = \mu_t + \frac{1}{2}\beta E_t bal_t. \quad (6)$$

Equations (1) - (3) also imply that

$$vol_{t+1} = -\alpha [(1 - \omega)spr_t + \omega E_t^* spr_{t+1} - \delta] + u_{t+1}^{vol}, \quad (7)$$

$$d_{t+1} = d_t - vol_{t+1} + u_{t+1}^d, \quad (8)$$

$$spd_t = \delta - \beta E_t d_t. \quad (9)$$

I assume that the common and directional components of market order and limit supply shocks:  $u_{t+1}^{flw}$ ,  $u_{t+1}^{bal}$ ,  $u_{t+1}^{vol}$  and  $u_{t+1}^d$  are *i.i.d.* mean zero random variables. Under these conditions, (4) - (9) show that the model exhibits a dichotomy between the behavior of prices, order flow and the depth balance on the one hand, and the behavior of the spread, trading volume and total depth on the other. Hereafter, I will refer to  $u_{t+1}^{flw}$ ,  $u_{t+1}^{bal}$ ,  $u_{t+1}^{vol}$  and  $u_{t+1}^d$  as exogenous shocks to order flow, balance, volume and depth, respectively.

Information plays a central role in the model. First, traders use information available at the start of period  $t$  to submit limit orders that determine  $p_t$  and  $spr d_t$  based on their estimates of fundamental value,  $\mu_t$ , the current depth balance,  $E_t bal_t$ , and total depth,  $E_t d_t$ . Since the actual depth on each side of the LOB is unobservable, these estimates are imprecise and are derived from past observations on prices and trade flows. The flows of market orders determining order flow and trading volume also depend on  $\mu_t$  and embed trader's expectations about future prices,  $E_t^* p_{t+1}$ , and spreads  $E_t^* spr_{t+1}$ . These expectations are conditioned on more information than was available at the start of period  $t$  when limit orders were determined. This assumption gives traders submitting market orders an informational advantage over those with outstanding limit orders.

Traders obtain information from outside and inside the LOB. Outside information comes at the start of each period in the form of news about the fundamental value of FX. Inside information comes in the form of period-by-period observations on prices, spreads, order flows and volume, and noisy signals concerning LOB depth. I now describe how these information sources are used to determine traders' expectations.

At the start of period  $t$ , traders revise their estimates of the fundamental value of FX according to

$$\mu_t = \mu_{t-1} + n_t + \lambda(flw_t - E_{t-1}flw_t), \quad (10)$$

with  $\lambda > 0$ . Here traders change their estimates between the start of periods  $t - 1$  and  $t$  based on news  $n_t$  which arrives at the start of period  $t$ , and the unexpected order flow during period  $t - 1$ . The news term  $n_t$  represents new information that directly leads to a revision in the estimated fundamental value of FX. The size of the  $\lambda$  parameter depends on the degree to which unexpected order flow is viewed as conveying new information about fundamental value.

After the arrival of news, traders determine limit orders based on their estimates of the depth balance  $bal_t$  and total depth  $d_t$ . These estimates depend on prior expectations and new information from market orders during period  $t - 1$ :

$$E_t bal_t = E_{t-1}^* bal_t + \varphi^{\text{bal}} (flw_t - E_{t-1}^* flw_t), \quad \text{and} \quad (11)$$

$$E_t d_t = E_{t-1}^* d_t - \varphi^d (vol_t - E_{t-1}^* vol_t), \quad (12)$$

where  $\varphi^{\text{bal}}$  and  $\varphi^d$  are positive constants less than one. According to (11), traders use their observations of order flow during period  $t - 1$  to revise their prior expectations about the depth balance (i.e.,  $E_{t-1}^* bal_t$ ). These flows are informative because (5) shows that changes in the depth balance between the start of periods  $t - 1$  and  $t$  depend on order flow during period  $t - 1$ . Similarly, in (12) traders use their observation on volume during period  $t - 1$  to revise their prior estimate of total depth at the start of period  $t$  (i.e.,  $E_{t-1}^* d_t$ ). Again, volume is informative because (7) shows that volume during period  $t - 1$  contributes to the change in total depth between the start of periods  $t - 1$  and  $t$ .

Next, traders receive noisy signals about the current depth balance and total depth in the LOB. The balance and depth signals are equal to  $bal_t + \xi_t^{\text{bal}}$  and  $d_t + \xi_t^d$ , respectively. Here  $\xi_t^{\text{bal}}$  and  $\xi_t^d$  are independently distributed mean-zero random variables that represent the noise in each signal. Expectations incorporate these signals as

$$E_t^* bal_t = E_t bal_t + \kappa^{\text{bal}} (bal_t + \xi_t^{\text{bal}} - E_t bal_t), \quad \text{and} \quad (13)$$

$$E_t^* d_t = E_t d_t + \kappa^d (d_t + \xi_t^d - E_t d_t), \quad (14)$$

where  $\kappa^{\text{bal}}$  and  $\kappa^d$  are positive constants less than one. Recall that traders make decisions about the submission of market orders during period  $t$  based on  $E_t^* bal_t$  and  $E_t^* d_t$  rather than  $E_t bal_t$  and  $E_t d_t$ . These decisions are therefore based on more precise information concerning the structure of the LOB than was available when limit orders were determined at the start of period  $t$ . As we shall see, this informational advantage plays an important role in the model.

This model can be viewed as an extension of earlier models where investors trade sequentially with an FX dealer. (see, e.g., Lyons, 1995 and 2001). In that setting, the dealer quotes prices based on estimated



fundamental value and their inventory. They revise their estimates of fundamental value based on news and order flow in a similar manner to (10), and these revisions are reflected one-for-one in the price level as in (6). Dealers also control their inventories by lowering (raising) their price quotes (relative to estimated fundamental value) when their inventory rises above (below) its targeted level. The estimated depth balance plays a similar role in this model. Here competition between traders submitting limit orders pushes the (mid-point) price level below (above) estimated fundamental value when the estimated depth of sell orders exceeds that of buy orders. The determination of market orders in this model can also be seen as an extension of the role played by investors in earlier models. There investors' trading decisions depend on the difference between the dealer's quotes and estimated fundamental value. Here market orders similarly depend on the difference between the trade price and estimated fundamental value, but the trade price must be estimated by individual traders because there is limited information about the matching of their order with existing limit orders.

It is also worth acknowledging that this is not a structural model in the sense that the equations for prices and trade flows in (4), (6), (7) and (9) are not explicitly derived from the optimal decisions of traders. Recent microstructure research considers the optimal placement of limit orders (see e.g., Glosten, 1994) and the choice between market and limit orders (see, e.g., Foucault, 1999), but deriving the equilibrium in a LOB based on fully optimal trading decisions remains a challenge (see, Parlour and Seppi, 2008 and Foucault et al., 2013 for overviews). I have deliberately chosen to avoid this challenge in order to provide a tractable framework for studying the transmission of different shocks into FX prices and trade flows when traders have limited information about the state of the LOB.

## 1.2 Order Flow and Price Dynamics

I now examine how FX prices and order flow react to different shocks. This analysis draws on the solution of the model which is derived in the Appendix. The equilibrium behavior of order flow and prices is most simply described by the dynamics of two variables: the estimated depth balance  $E_t bal_t$  and the associated estimation error  $er_t^{bal} = bal_t - E_t bal_t$ . These variables follow a stationary vector process:

$$\begin{bmatrix} E_t bal_t \\ er_t^{bal} \end{bmatrix} = \begin{bmatrix} \rho & (1 - \omega\phi)\kappa^{bal} \\ 0 & 1 - \kappa^{bal} \end{bmatrix} \begin{bmatrix} E_{t-1} bal_{t-1} \\ er_{t-1}^{bal} \end{bmatrix} + \begin{bmatrix} \varphi^{bal} & 0 & (1 - \omega\phi)\kappa^{bal} \\ 1 - \varphi^{bal} & 1 & -\kappa^{bal} \end{bmatrix} \begin{bmatrix} u_t^{flw} \\ u_t^{bal} \\ \xi_{t-1}^{bal} \end{bmatrix} \quad (15)$$

where  $\rho = \frac{1}{1+\alpha\beta}$  and  $\phi = \frac{\alpha\beta}{1+\alpha\omega(2\lambda+\beta)}$ . Order flow during period  $t$  depends on the estimated depth balance and its associated estimation error:

$$flw_{t+1} = -(1 - \rho)E_t bal_t + u_{t+1}^{flw} - \omega\phi\kappa^{bal}(er_t^{bal} + \xi_t^{bal}). \quad (16)$$

The last two terms are unforecastable at the start of period  $t$ , so the revision in estimated fundamental value between  $t$  and  $t + 1$  is given by

$$\mu_{t+1} = \mu_t + n_{t+1} + \lambda [u_{t+1}^{\text{flw}} - \omega\phi\kappa^{\text{bal}}(er_t^{\text{bal}} + \xi_t^{\text{bal}})]. \quad (17)$$

Combining (15) and (17) with (6) gives the following equation for the change in prices:

$$\begin{aligned} \Delta p_{t+1} = & n_{t+1} + \left(\lambda + \frac{1}{2}\beta\varphi^{\text{bal}}\right) u_{t+1}^{\text{flw}} \\ & - \frac{1}{2}\beta(1 - \rho)E_t bal_t + \frac{1}{2} [\beta - \omega\phi(\beta + 2\lambda)] \kappa^{\text{bal}}(er_t^{\text{bal}} + \xi_t^{\text{bal}}). \end{aligned} \quad (18)$$

These equations display several noteworthy features. First, since the estimated balance  $E_t bal_t$  contributes to order flow and price changes, both variables display persistence and can be forecast using information available at the start of period  $t$ . Second, incomplete information concerning the state of the LOB contributes to the persistence of order flow and price changes. Equation (15) shows that the estimation errors  $er_t^{\text{bal}}$  are serially correlated and contribute to the dynamics of  $E_t bal_t$ . Third, noisy observations on the depth balance provide traders with more precise information about order flow and future price changes when choosing their market orders than was available when limit orders were determined. The last terms on the right-hand-side of equations (16) and (18) identify the new information conveyed by the noisy observations so that

$$\begin{aligned} (E_t^* - E_t)flw_{t+1} &= -\omega\phi\kappa^{\text{bal}}(er_t^{\text{bal}} + \xi_t^{\text{bal}}) \quad \text{and} \\ (E_t^* - E_t)\Delta p_{t+1} &= \frac{1}{2} [\beta - \omega\phi(\beta + 2\lambda)] \kappa^{\text{bal}}(er_t^{\text{bal}} + \xi_t^{\text{bal}}). \end{aligned}$$

These features come into play as we next examine the effects of different shocks.

The transmission of news shocks occurs entirely via the information channel. To see why, consider the effects of a positive news shock  $n_{t+1}$  that arrives at the start of period  $t + 1$ . Equation (17) shows that the arrival of news leads to an immediate upward revision in the estimated fundamental value of FX. This revision is reflected in period- $t + 1$  limit orders, increasing both the best bid and ask prices equally, so there is a one-to-one price change, as shown in equation (18). Moreover, because the news produces no further revisions in estimated fundamental value beyond  $t + 1$ , it has no effect on the expected future price changes that drive market orders during period  $t + 1$  and beyond. Consequently, as (16) shows, news shocks have no effect on order flow. This feature of the model is entirely standard. It reflects the fact that news shocks revise the estimates of fundamental value used by all traders submitting limit and market orders.

Order flow shocks affect prices and order flows through two channels: the liquidity and information channels. The liquidity channel operates by changing traders' estimates of the depth balance. Equation (16) shows that a positive  $u_{t+1}^{\text{flw}}$  shock produces an unexpected rise in period- $t$  order flow  $flw_{t+1}$ , which in turn produces an increase in the estimated balance  $E_{t+1} bal_{t+1}$ , as shown in (15). Thereafter, the effects of the shock dissipate until estimates of the depth balance return to zero. Consequently, since  $p_{t+1} = \mu_{t+1} + \frac{1}{2}\beta E_{t+1} bal_{t+1}$ , the positive  $u_{t+1}^{\text{flw}}$  shock pushes prices above the estimated fundamental value of FX from period

$t + 1$  until their effects on the estimated depth balances disappears. The information channel operates by changing the estimated fundamental value of FX. In particular, (17) shows that a positive  $u_{t+1}^{\text{flw}}$  shock induces an upward revision in estimated fundamental value at the start of period  $t + 1$  equal to  $\lambda u_{t+1}^{\text{bal}}$ . Thereafter, the shock adds to the estimation errors  $er_{t+i}^{\text{bal}}$  for  $i > 1$  as shown in (15). These errors lead to further revisions in estimated fundamentals because they contribute to future unexpected order flows. Equation (17) implies that the cumulative effect of a positive  $u_{t+1}^{\text{flw}}$  shock is to increase estimated fundamental value by  $[\lambda - \omega\phi(1 - \varphi^{\text{bal}})]u_{t+1}^{\text{bal}}$ . Since changes in  $\mu_t$  are reflected one-to-one in the price level, the informational channel gives order flow shocks a smaller permanent price impact than their immediate price impact. Equation (18) implies that the immediate price impact of order flow shocks operating through the liquidity and information channels is  $\lambda + \frac{1}{2}\beta\varphi^{\text{bal}}$ , while the permanent price impact is  $\lambda + \frac{1}{2}\beta - \omega\phi(\beta + 2\lambda)(1 - \varphi^{\text{bal}})$ . For reasonable parameterizations of the model, the permanent price impact is slightly smaller than  $\lambda$ , so prices fall after period  $t + 1$  until they reach their new long-run level. During this period, order flows are negative because traders estimate that trade prices are above the estimated fundamental value of FX.

The operation of the liquidity and information channels discussed above is affected by the presence of limited information. In particular, limited information adds to the persistence of the liquidity channel because order flow shocks have longer lasting effects on the estimated depth balance.<sup>8</sup> The liquidity effects of order flow shocks are still present when traders know the depth balance, and they still temporarily push prices away from estimated fundamental value, but the effect dies out more quickly. Limited information also contributes to the persistence of the information channel. If traders know the depth balance directly, all the information effects of the order flow shock occur immediately because the  $u_{t+1}^{\text{flw}}$  shock only contributes to unexpected order flows during period  $t$ . Consequently, order flows have an immediate and permanent price impact through the information channel.

Next, consider the effects of a positive balance shock  $u_{t+1}^{\text{bal}}$  that increases the depth balance (i.e. the depth of limit buy orders relative to sell orders) at the start of period  $t + 1$ . This shock has no immediate effect on  $p_{t+1}$  because traders setting limit orders at the start of period  $t + 1$  are unaware of the shock. In contrast, the shock produces an unexpected fall in period  $t + 1$  order flow because traders revise their estimates for trade prices upwards based on the higher signal concerning  $bal_{t+1}$  they receive before placing their market orders. From this point forward, the effects of the shock are transmitted to prices via the liquidity and information channels. Traders submitting limit orders use their higher estimates of the depth balance to push prices above fundamental value. They also revise their estimates of fundamental value downwards in response to the unexpected fall in order flow. Thus, the information and liquidity channels initially push prices in opposite directions: the liquidity channel pushes the price upward, while the information channel pulls it downwards. This means that while positive balance shocks induce an unexpected fall in order flow, they can produce either an unexpected rise or fall in prices depending upon whether the liquidity or information channels dominate. Of course, as time passes the liquidity effects of the shock die out, so that the permanent price impact of a positive balance shock is negative.<sup>9</sup>

<sup>8</sup>To see this analytically, note that (15) simplifies to  $E_t bal_t = \rho E_{t-1} bal_{t-1} + u_t^{\text{bal}} + u_t^{\text{flw}}$  when traders observe the depth balance directly ( $bal_t = E_t bal_t$ ).

<sup>9</sup>Traders' misperceptions about the state of the LOB can also affect prices and order flow. In particular, noise in signals

Two key features of the model give rise to these price and order flow responses: (i) the informational advantage of traders submitting market orders over limit orders, and (ii) the limited depth at the top of the LOB. If traders submitted market orders before receiving the signal about the depth balance (thereby eliminating their information advantage over traders submitting limit orders), the the balance shock would not produce unexpected order flow in period  $t + 1$ . In fact, it would only contribute to lower *expected* order flow in the periods that follow. Under these circumstances, the shock would have no effect on prices via the information channel, it would only temporarily raise prices via the liquidity channel. Similarly, if there was unlimited depth on either side of the LOB at the best prices, traders submitting market orders would know that their orders would be executed at those prices. The model accommodates this special case when  $\omega = 0$ . Under these circumstances, information about the balance shock would not impact on the traders' choice of market orders, so again order flow in period  $t + 1$  would be unaffected and the information channel would be inoperable.

To better understand the roles of the information and liquidity channels, Figure 1 plots the responses of the (log) price level and order flow to order flow shocks and balance shocks. The plots are computed from a simulation of the model with  $\alpha = 15$ ,  $\beta = 0.1$ ,  $\lambda = 0.2$ ,  $\kappa^{\text{bal}} = \kappa^{\text{d}} = 0.6$  and  $\varphi^{\text{bal}} = \varphi^{\text{d}} = 0.5$ : the implied value for  $\rho$  is 0.86.

Panels A and B show the responses to a positive order flow shock  $u_{t+1}^{\text{fw}}$  that occurs during period  $t$ . Here we see how the effects of the shock operating through the information and liquidity channels lead to an immediate price-impact (in period  $t + 1$ ) that exceeds the long-run impact (depicted by the dashed line in panel A). Panel B shows that the shock induces a persistent effect on order flow. After the initial effect of the shock, the liquidity channel temporarily pushes order flow negative while prices are above the estimated fundamental value of FX.

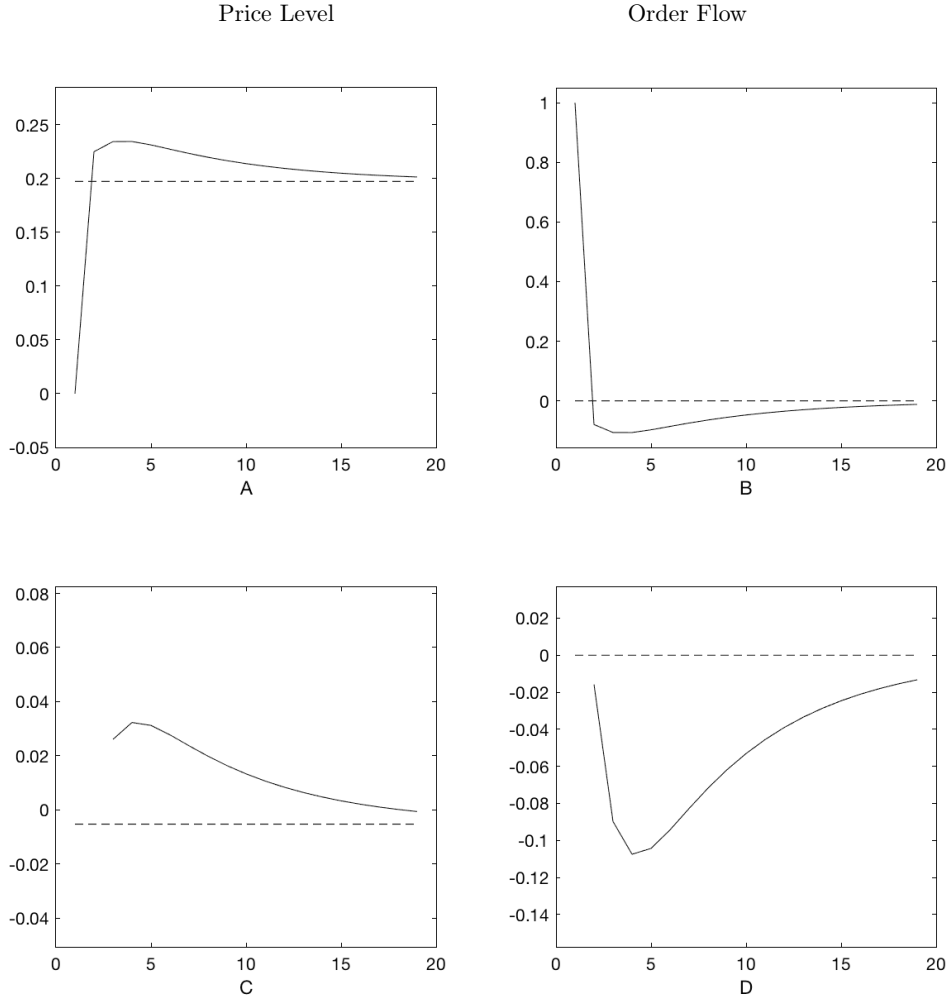
The effects of a positive period  $t$  balance shock  $u_t^{\text{bal}}$  are shown in panels C and D. In this calibration of the model, the liquidity channel is stronger than the information channel so the immediate price-impact of the shock is positive. Of course, over time, the liquidity effects dissipate so that prices fall towards the new estimate of fundamental value, shown by the dashed line in panel C. In panel D we see that the shock initially produces an unexpected fall in order flow. This is then followed by a period of negative order flow while prices remain above estimated fundamental value.

The plots in Figure 1 make clear that the contemporaneous correlation between unexpected price changes and order flow depends on the source of the shock. In this particular calibration of the model exogenous shocks to order flow produce a positive correlation, whereas balance shocks produce a negative correlation. These correlations reflect the relative importance of the information and liquidity channels in the transmission of different shocks to prices. The plots in Figure 1 also show that the long-run price impact of a shock can differ from its immediate impact when liquidity effects are important. The long run price impact of order flow shocks is smaller because the liquidity effects amplify the immediate information effects of the shock. In contrast, the liquidity effects completely dominate the immediate information effects of balance shocks,

---

traders receive about the depth balance affect order flows and prices through the information and liquidity channels in a similar manner as balance shocks.

**Figure 1: Model Impulse Responses**



Notes: Impulse responses of the (log) price level, and order flow to positive order flow shocks  $u_t^{\text{flw}}$  (panels A and B), and balance shocks  $u_t^{\text{bal}}$  (panels C and D). Dashed lines show long run response.

so that the short- and long-run price impacts are in opposite directions.

In summary, the trading model illustrates the complex process by which shocks affect FX prices and trading flows in a LOB. The complexity arises because the shocks operate through multiple channels that have different short- and long-run effects. In the next section, I empirically examine this complex process.

## 2 Intraday Empirical Analysis

I now describe the structural VAR model I estimate with LOB data on trading in the EURUSD, the world's most heavily traded currency pair. The trading model in Section 1 includes several stylized but counterfactual assumptions that facilitated the analysis of equilibrium dynamics. In particular, the model assumes that the submission of limit and market orders follow a sequence, and that noisy information about the depth of

the LOB reaches traders in a specific way. In reality, the submission of limit and market orders can occur simultaneously and continuously. Furthermore, while actual traders have limited information about the state of the LOB, they undoubtedly receive new information in a more complex manner than is assumed in the trading model. In light of these observations, I will not attempt to empirically estimate the exact structure of the trading model. Instead, I use its implications concerning the short- and long-run effects of shocks to impose restrictions on the VAR so that the estimates are informative about the sources of FX price dynamics and trading flows.

## 2.1 The Structural VAR

The VAR includes six variables from the trading model: order flow and volume during period  $t$ ,  $flw_{t+1}$  and  $vol_{t+1}$ ; the depth balance, total depth, and the spread at the start of period  $t + 1$ ,  $bal_{t+1}$ ,  $d_{t+1}$  and  $spd_{t+1}$ ; and the change in (log) mid-point prices between the start of periods  $t$  and  $t + 1$ ,  $\Delta p_{t+1} = p_{t+1} - p_t$ . One period in the VAR spans 30 seconds. While this is clearly a much shorter time span than is typically used in VARs, it often covers a great deal of EURUSD trading activity. Examining the LOB trading data at a 30-second frequency provides greater clarity into the process driving prices and trade flows than is possible at lower frequencies (e.g., five minutes) while avoiding the technical complexity of estimating a model on a trade-by-trade basis (see, e.g. Evans, 2011).

I assume that the vector  $Y_{t+1} = [flw_{t+1}, vol_{t+1}, bal_{t+1}, d_{t+1}, spd_{t+1}, \Delta p_{t+1}]'$  follows a (possibly infinite) covariance stationary vector moving average process:

$$Y_{t+1} = \Theta(L)V_{t+1}, \quad (19)$$

where  $\Theta(L) = \Theta_0 + \Theta_1 L + \Theta_2 L^2 + \dots$  is a matrix polynomial in the lag operator and  $V_{t+1} = [v_{t+1}^1, \dots, v_{t+1}^6]'$  is a  $6 \times 1$  vector of mean-zero innovations with diagonal covariance matrix  $\Sigma_V$ .

I place two sets of restrictions on the matrices of moving average coefficients: short-run restrictions on the impact matrix  $\Theta_0$ , and long-run restrictions on the sum of the  $\Theta_i$  matrices,  $\Psi = \sum_i \Theta_i$ . The restrictions are

$$\Theta_0 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \theta_{16} \\ \theta_{21} & 1 & 0 & 0 & 0 & 0 \\ \theta_{31} & \theta_{32} & 1 & 0 & 0 & \theta_{36} \\ \theta_{41} & \theta_{42} & \theta_{43} & 1 & \theta_{45} & 0 \\ \theta_{51} & \theta_{52} & \theta_{53} & \theta_{54} & 1 & 0 \\ \theta_{61} & \theta_{62} & \theta_{63} & \theta_{64} & \theta_{65} & 1 \end{bmatrix} \quad \text{and} \quad \Psi = \begin{bmatrix} \psi_{11} & \psi_{12} & \psi_{13} & \psi_{14} & \psi_{15} & \psi_{16} \\ \psi_{21} & \psi_{22} & \psi_{23} & \psi_{24} & \psi_{25} & \psi_{26} \\ \psi_{31} & \psi_{32} & \psi_{33} & \psi_{34} & \psi_{35} & \psi_{36} \\ \psi_{41} & \psi_{42} & \psi_{43} & \psi_{44} & \psi_{45} & \psi_{46} \\ \psi_{51} & \psi_{52} & \psi_{53} & \psi_{54} & \psi_{55} & \psi_{56} \\ \psi_{61} & \psi_{62} & 0 & 0 & 0 & \psi_{66} \end{bmatrix}. \quad (20)$$

The impact matrix  $\Theta_0$  contains 12 zeros that restrict the contemporaneous impact of innovations on VAR variables. The coefficients on the leading diagonal of  $\Theta_0$  are normalized to one to facilitate interpretation of the model estimates. The long-run matrix  $\Psi$  contains just three zero restrictions in the sixth row. Because the change in FX prices appears as the sixth variable in  $Y_{t+1}$ , these three zeros restrict the third, fourth and

fifth innovations to have no permanent effect on the level of prices. The 18  $\theta_{i,j}$  coefficients and six innovation variances can be exactly identified from estimates of the VAR by imposing the long-run restrictions in  $\Psi$ .

The restrictions on the impact matrix are conventional except for the three (possibly) non-zero coefficients that appear above the leading diagonal. The motivation for these coefficients comes from the trading model. Recall that the contemporaneous correlation between unexpected price changes and order flow could be positive, negative or zero depending on the shock. Here I allow the correlation between price changes and order flow to be similarly determined by the source of the innovation. In particular, the first innovation  $v_t^1$  has a one-for-one impact on order flow  $flw_t$  while inducing a price change  $\Delta p_{t+1}$  of  $\theta_{61}v_t^1$ . From the perspective of the trading model, it is natural to interpret  $v_t^1$  as an order flow shock, so we would expect to estimate a positive value for  $\theta_{61}$  (implying a positive correlation between price changes and order flow).<sup>10</sup> The structure of the  $\Theta_0$  matrix also allows  $v_t^1$  innovations to affect the other four variables in the model.<sup>11</sup> The sixth innovation,  $v_t^6$ , can induce a different correlation between price changes and order flow. According to the impact matrix,  $v_t^6$  has a one-to-one effect on prices while inducing an order flow equal effect of  $\theta_{16}v_t^6$ . If news shocks are the primary contributors to the  $v_t^6$  innovations,  $\theta_{16}$  should be (close to) zero because news shocks have no effects on order flow (implying a zero correlation between price changes and order flow). In this case,  $\theta_{36}$  should also be (close to) zero because news shocks have no effect on the depth balance. Alternatively, if  $v_t^6$  innovations are primarily driven by balance shocks as in the trading model,  $\theta_{16}$  should be negative.

The structure of the  $\Theta_0$  matrix also allows the correlation between the spread and depth to depend on the source of the innovation. In the trading model, changes in the spread are negatively correlated with revisions in the estimated depth of the LOB. So shocks to actual depth that are accurately perceived produce a negative correlation between actual depth and the spread. In contrast, shocks that inaccurately induce a revision in estimated depth induce a movement in the spread without a corresponding change in actual depth. The coefficients  $\theta_{54}$  and  $\theta_{45}$  allow for similar differences in the correlation between depth and spreads depending on the incidence of the fourth and fifth innovations,  $v_t^4$  and  $v_t^5$ .

The structure of the  $\Psi$  matrix allows for just three of innovations,  $v_t^1$ ,  $v_t^2$  and  $v_t^6$ , to have a long-run effect on the price level. In the trading model, shocks have a permanent price-impact via the information channel. So from that perspective,  $v_t^1$ ,  $v_t^3$  and  $v_t^6$  will have long-run price effects insofar as they reflect shocks that affect prices via the information channel. For example, if  $v_t^1$  innovations reflect order flow shocks, we should expect  $\psi_{61}$  to be positive, but less than  $\theta_{61}$  because the liquidity channel amplifies the short-run effects of order flow shocks on prices. Similarly, if news shocks are the primary source of  $v_t^6$  innovations,  $\psi_{66}$  should equal one because the price-impact of such shocks is immediate and complete (i.e., the short- and long-run price effects are equal). Alternatively, if balance shocks are the main contributor to the  $v_t^6$  innovations, we expect the long-run price effects to differ from the short-run effects, as illustrated in Figure 1. In this case,  $\psi_{66}$  could be positive or negative. The structure of the  $\Theta_0$  matrix also allows for the *possibility* that  $v_t^2$

<sup>10</sup>While this is a natural *interpretation*, the VAR innovations are logically distinct from (linear combinations of) the exogenous shocks in the trading model because the linear VAR forecasts for  $Y_{t+1}$  need not equal traders' conditional expectation,  $E_t Y_{t+1}$ .

<sup>11</sup>Under the interpretation that  $v_t^1$  represent order flow shocks,  $\theta_{31}$  should be positive, while  $\theta_{21}$ ,  $\theta_{41}$  and  $\theta_{51}$  should be close to zero. These predictions are borne out in the estimates presented below.

innovations have a long-run price impact identified by the  $\psi_{62}$  coefficient. This could occur if order flow or balance shocks contribute to the  $v_t^2$  innovations.

In summary, the structural VAR imposes restrictions that can accommodate the various correlations between order flow and price changes produced by different shocks, and differences between their immediate and permanent price-impacts. Both of these features turn out to be empirically important.

## 2.2 Data

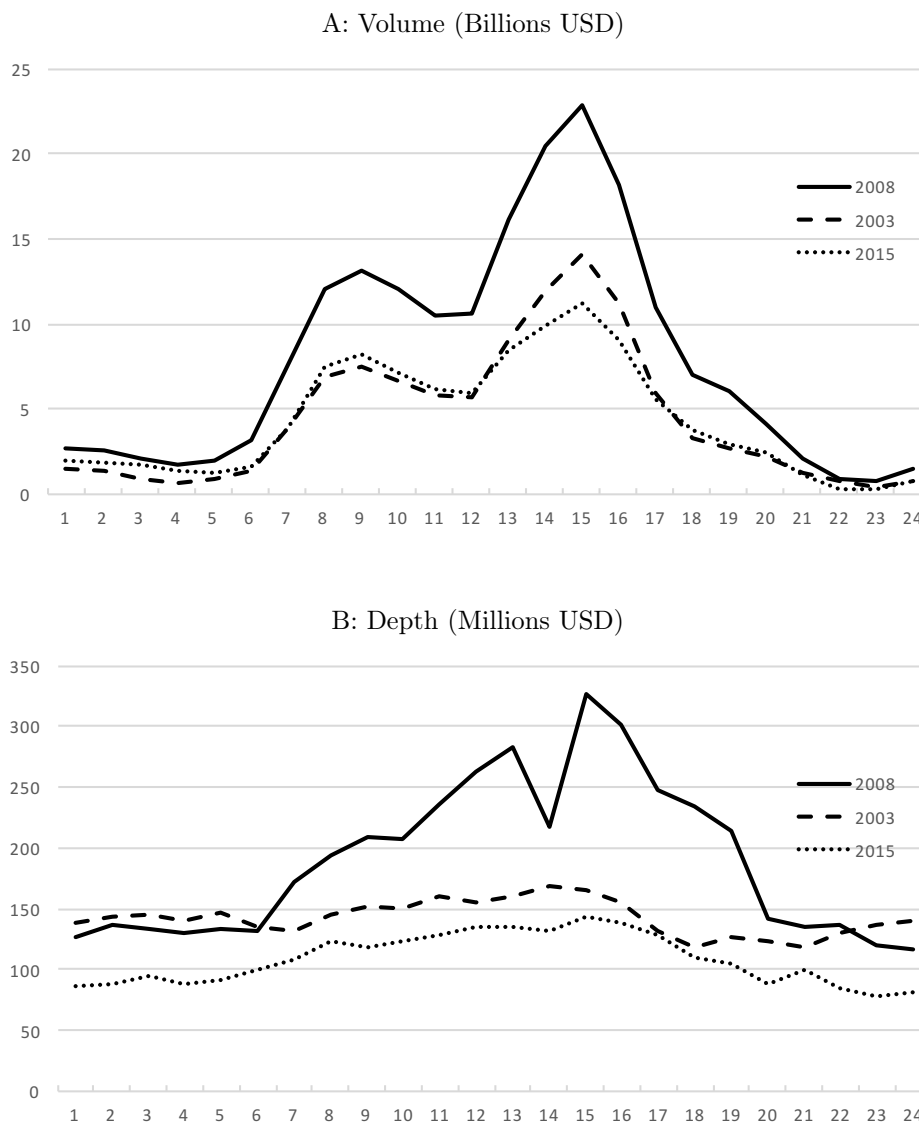
I construct the six VAR variables from the EBS trading data using an observation window of 30 seconds. Although FX trading takes place continuously 24 hours a day, seven days a week, there are distinct weekly and daily patterns in activity that show up in volume, depth and the spread. To illustrate these seasonal patterns, Figure 2 plots median volume and median depth for each hour in three years from the sample period. The plots in panel A show that trading volume is typically concentrated between 7:00 and 18:00 hrs (London time), with a distinct peak between 14:00 and 16:00 hrs. Notice also that trading volume is significantly higher in 2008 than in 2003 and 2015. These differences in volume are representative of a persistent increase in volume between 2003 and 2010, followed by a decline from 2011 to 2015. Panel B shows that the median depth of the LOB showed much less intraday “seasonality” than volume during 2003 and 2015. In contrast, median depth increased markedly between 7:00 and 18:00 hrs during 2008, peaking at approximately the same time as volume.

I remove these weekly and daily patterns before estimating the VAR. First, I exclude the 48 hours between midnight (London Time) on Friday and Sunday because there is very little activity during this period. Second, I estimate the average volume, depth and spread for each 15-minute interval in the remaining days using data from the same interval on the prior 10 trading days. For example, I estimate seasonal volume between 12:00 and 12:15 pm on day 11 as the average volume (per 30-second period) between 12:00 and 12:15 pm on days 1 to 10. This procedure only uses information about past daily patterns of activity. It also accommodates the slow-moving change in the seasonal patterns that occurred over the months and years spanned by the sample that was illustrated by Figure 2.

Table 1 provides statistics for the six variables in the VAR. The variables are defined as follows: Order flow  $flw_{t+1}$  is computed as the difference between the value of market EUR purchase and market EUR sale orders in millions of USD during the 30-second period  $t$ . Volume  $vol_{t+1}$  is the sum market EUR purchase and market EUR sale orders in millions of USD during period  $t$ . The depth balance  $bal_t$  is measured as the difference in millions between the USD value of limit bids to buy EUR and limit offers to sell EUR at the start of period  $t$ . Total depth  $d_t$  is the USD value in millions of limit offers to sell EUR and limit bids to buy EUR at the start of period  $t$ . The spread and price change variables (measured in basis points) are computed from the logs of the best limit bid and offer prices,  $p_t^{\text{bid}}$  and  $p_t^{\text{offer}}$ , at the start of period  $t$ . In particular, the period- $t$  spread is computed as  $spd_t = (p_t^{\text{offer}} - p_t^{\text{bid}}) \times 10000$ , and the price change between  $t$  and  $t + 1$  is computed as  $\Delta p_{t+1} = (p_{t+1} - p_t) \times 10000$ , where  $p_t = \frac{1}{2}p_t^{\text{offer}} + \frac{1}{2}p_t^{\text{bid}}$ .



**Figure 2: Seasonal Volume and Depth**



Notes: Panel A plots the median trading volume in billions for each hour (London Time) for three years from the sample period: 2003, 2008 and 2015. Panel B plots the median depth in millions for each hour. Trading data between midnight on Friday and midnight on Sunday are excluded from the calculations.

Panel A of Table 1 reports the mean, median, standard deviation and autocorrelations for the raw variables. The most notable statistics are the autocorrelations, which cover up to 20 periods, or equivalently 10 minutes of trading activity. Volume, total depth, and the depth balance are all strongly autocorrelated, whereas order flow, the spread and price changes display small amounts of serial correlation. Seasonal patterns contribute to the autocorrelations in some of these variables.<sup>12</sup> Panel B reports statistics for “de-

<sup>12</sup>There is no discernible seasonal pattern in the depth balance.

seasonalized” volume, depth and the spread. These variables are computed by removing their respective seasonal components (estimated from the prior 10 trading days as described above). De-seasonalization reduces the size of autocorrelations more at longer lags than shorter lags. It also reduces the standard deviation of the volume and depth.

**Table 1: Summary Statistics**

	Order Flow	Volume	Depth Balance	Total Depth	Spread	Price Change	
A:							
Mean	0.136	50.719	0.626	177.663	1.048	0.000	
Median	0.000	24.000	1.000	148.000	0.824	0.000	
Std.	26.690	74.864	79.673	123.698	1.967	1.311	
Autocorr.							
lag =	1	0.101	0.753	0.908	0.963	0.053	-0.036
	5	0.025	0.608	0.811	0.925	0.031	-0.004
	10	0.017	0.566	0.747	0.900	0.029	0.001
	20	0.013	0.523	0.665	0.868	0.026	0.001
B:							
Mean		-3.307		-0.692	0.000		
Median		-8.829		-8.325	-0.086		
Std.		60.363		80.075	1.945		
Autocorr.							
lag =	1		0.620		0.911	0.037	
	5		0.398		0.817	0.016	
	10		0.333		0.755	0.014	
	20		0.266		0.676	0.010	

Notes: The table reports summary statistics for the variables in the VAR. Panel A shows statistics for the raw variables, while panel B shows statistics for the de-seasonalized variables. All statistics are computed from 9760079 observations, at a 30-second frequency, over 13 years; January 6, 2003 to December 31, 2015.

## 2.3 VAR Estimates

Table 2 shows estimates of the impact matrix  $\Theta_0$  and the innovation variances in the covariance matrix  $\Sigma_V$ . These estimates are computed from a third-order VAR; estimates from higher-order VARs are very similar. The table also shows standard errors in parentheses under each of the estimated parameters. These standard errors (and the confidence bands shown below) are computed from a bootstrap with 5000 replications. The VAR estimates and the bootstrap are computed from time series with approximately 9.8 million observations, so the standard errors associated with many of the parameters are extremely small. The table shows a value of 0.0000 in the cases where the bootstrap standard errors are smaller than 0.0001. Restricted coefficients in the impact matrix (i.e. those equal to one or zero) appear without standard errors. By conventional standards, almost all the coefficient estimates in the impact matrix are highly statistically significant.

Our primary interest is in understanding the origins of movements in order flows and FX prices. To this

**Table 2: VAR Estimates**

$\Theta_0 =$	<table style="border: none; border-collapse: collapse;"> <tr><td style="border: none; padding-right: 10px;">1.0000</td><td style="border: none; padding-right: 10px;">0.0000</td><td style="border: none; padding-right: 10px;">0.0000</td><td style="border: none; padding-right: 10px;">0.0000</td><td style="border: none; padding-right: 10px;">0.0000</td><td style="border: none; padding-right: 10px;">0.0000</td><td style="border: none; padding-right: 10px;">-4.1103</td></tr> <tr><td style="border: none; padding-right: 10px;">(0.0008)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0018)</td></tr> <tr><td style="border: none; padding-right: 10px;">-0.0134</td><td style="border: none; padding-right: 10px;">1.0000</td><td style="border: none; padding-right: 10px;">0.0000</td><td style="border: none; padding-right: 10px;">0.0000</td><td style="border: none; padding-right: 10px;">0.0000</td><td style="border: none; padding-right: 10px;">0.0000</td><td style="border: none; padding-right: 10px;">0.0000</td></tr> <tr><td style="border: none; padding-right: 10px;">(0.0005)</td><td style="border: none; padding-right: 10px;">(0.0002)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0027)</td></tr> <tr><td style="border: none; padding-right: 10px;">0.0611</td><td style="border: none; padding-right: 10px;">-0.0009</td><td style="border: none; padding-right: 10px;">1.0000</td><td style="border: none; padding-right: 10px;">0.0000</td><td style="border: none; padding-right: 10px;">0.0000</td><td style="border: none; padding-right: 10px;">0.0000</td><td style="border: none; padding-right: 10px;">-0.1943</td></tr> <tr><td style="border: none; padding-right: 10px;">(0.0005)</td><td style="border: none; padding-right: 10px;">(0.0002)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0027)</td></tr> <tr><td style="border: none; padding-right: 10px;">-0.0076</td><td style="border: none; padding-right: 10px;">-0.0263</td><td style="border: none; padding-right: 10px;">-0.0127</td><td style="border: none; padding-right: 10px;">1.0000</td><td style="border: none; padding-right: 10px;">-1.5245</td><td style="border: none; padding-right: 10px;">0.0000</td><td style="border: none; padding-right: 10px;">0.0000</td></tr> <tr><td style="border: none; padding-right: 10px;">(0.0005)</td><td style="border: none; padding-right: 10px;">(0.0002)</td><td style="border: none; padding-right: 10px;">(0.0004)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0023)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0000)</td></tr> <tr><td style="border: none; padding-right: 10px;">0.0013</td><td style="border: none; padding-right: 10px;">0.0023</td><td style="border: none; padding-right: 10px;">-0.0001</td><td style="border: none; padding-right: 10px;">0.0420</td><td style="border: none; padding-right: 10px;">1.0000</td><td style="border: none; padding-right: 10px;">0.0000</td><td style="border: none; padding-right: 10px;">0.0000</td></tr> <tr><td style="border: none; padding-right: 10px;">(0.0001)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0001)</td><td style="border: none; padding-right: 10px;">(0.0001)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0000)</td></tr> <tr><td style="border: none; padding-right: 10px;">0.5274</td><td style="border: none; padding-right: 10px;">-0.0008</td><td style="border: none; padding-right: 10px;">-0.0393</td><td style="border: none; padding-right: 10px;">-0.0018</td><td style="border: none; padding-right: 10px;">0.0139</td><td style="border: none; padding-right: 10px;">1.0000</td><td style="border: none; padding-right: 10px;">0.0000</td></tr> <tr><td style="border: none; padding-right: 10px;">(0.0001)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0003)</td><td style="border: none; padding-right: 10px;">(0.0000)</td><td style="border: none; padding-right: 10px;">(0.0000)</td></tr> </table>	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-4.1103	(0.0008)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0018)	-0.0134	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	(0.0005)	(0.0002)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0027)	0.0611	-0.0009	1.0000	0.0000	0.0000	0.0000	-0.1943	(0.0005)	(0.0002)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0027)	-0.0076	-0.0263	-0.0127	1.0000	-1.5245	0.0000	0.0000	(0.0005)	(0.0002)	(0.0004)	(0.0000)	(0.0023)	(0.0000)	(0.0000)	0.0013	0.0023	-0.0001	0.0420	1.0000	0.0000	0.0000	(0.0001)	(0.0000)	(0.0001)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	0.5274	-0.0008	-0.0393	-0.0018	0.0139	1.0000	0.0000	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0003)	(0.0000)	(0.0000)	$diag(\Sigma_V) =$	<table style="border: none; border-collapse: collapse;"> <tr><td style="border: none; padding-right: 10px;">5.2910</td></tr> <tr><td style="border: none; padding-right: 10px;">(0.0000)</td></tr> <tr><td style="border: none; padding-right: 10px;">25.7780</td></tr> <tr><td style="border: none; padding-right: 10px;">(0.0000)</td></tr> <tr><td style="border: none; padding-right: 10px;">11.2370</td></tr> <tr><td style="border: none; padding-right: 10px;">(0.0000)</td></tr> <tr><td style="border: none; padding-right: 10px;">10.5390</td></tr> <tr><td style="border: none; padding-right: 10px;">(0.0000)</td></tr> <tr><td style="border: none; padding-right: 10px;">(0.2443)</td></tr> <tr><td style="border: none; padding-right: 10px;">(0.0000)</td></tr> <tr><td style="border: none; padding-right: 10px;">0.2010</td></tr> <tr><td style="border: none; padding-right: 10px;">(0.0000)</td></tr> </table>	5.2910	(0.0000)	25.7780	(0.0000)	11.2370	(0.0000)	10.5390	(0.0000)	(0.2443)	(0.0000)	0.2010	(0.0000)
1.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-4.1103																																																																																													
(0.0008)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0018)																																																																																													
-0.0134	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000																																																																																													
(0.0005)	(0.0002)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0027)																																																																																													
0.0611	-0.0009	1.0000	0.0000	0.0000	0.0000	-0.1943																																																																																													
(0.0005)	(0.0002)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0027)																																																																																													
-0.0076	-0.0263	-0.0127	1.0000	-1.5245	0.0000	0.0000																																																																																													
(0.0005)	(0.0002)	(0.0004)	(0.0000)	(0.0023)	(0.0000)	(0.0000)																																																																																													
0.0013	0.0023	-0.0001	0.0420	1.0000	0.0000	0.0000																																																																																													
(0.0001)	(0.0000)	(0.0001)	(0.0001)	(0.0000)	(0.0000)	(0.0000)																																																																																													
0.5274	-0.0008	-0.0393	-0.0018	0.0139	1.0000	0.0000																																																																																													
(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0003)	(0.0000)	(0.0000)																																																																																													
5.2910																																																																																																			
(0.0000)																																																																																																			
25.7780																																																																																																			
(0.0000)																																																																																																			
11.2370																																																																																																			
(0.0000)																																																																																																			
10.5390																																																																																																			
(0.0000)																																																																																																			
(0.2443)																																																																																																			
(0.0000)																																																																																																			
0.2010																																																																																																			
(0.0000)																																																																																																			

Notes: The table reports estimates of the impact matrix  $\Theta_0$  and the innovation variances (from the leading diagonal of  $\Sigma_v$ ) computed from a third-order VAR estimated at a 30-second frequency, over 13 years; January 6, 2003 to December 31, 2015. Bootstrap standard errors are reported below each parameter estimate. Standard errors smaller than 0.0001 are shown as 0.0000.

end, Figure 3 plots the impulse responses of order flow and the log price *level* to the  $v_t^1$  and  $v_t^6$  innovations. The horizontal axis shows the number of minutes since the innovation. Table 3 shows below that these two innovations account for almost all the variance of order flows and price changes across all horizons, so I concentrate on these impulse responses.<sup>13</sup> In addition to the impulse responses, Figure 3 also shows 95% confidence intervals computed from the bootstrap as grey bands. The bands are almost invisible in the order flow responses because the underlying VAR parameters are so precisely estimated.

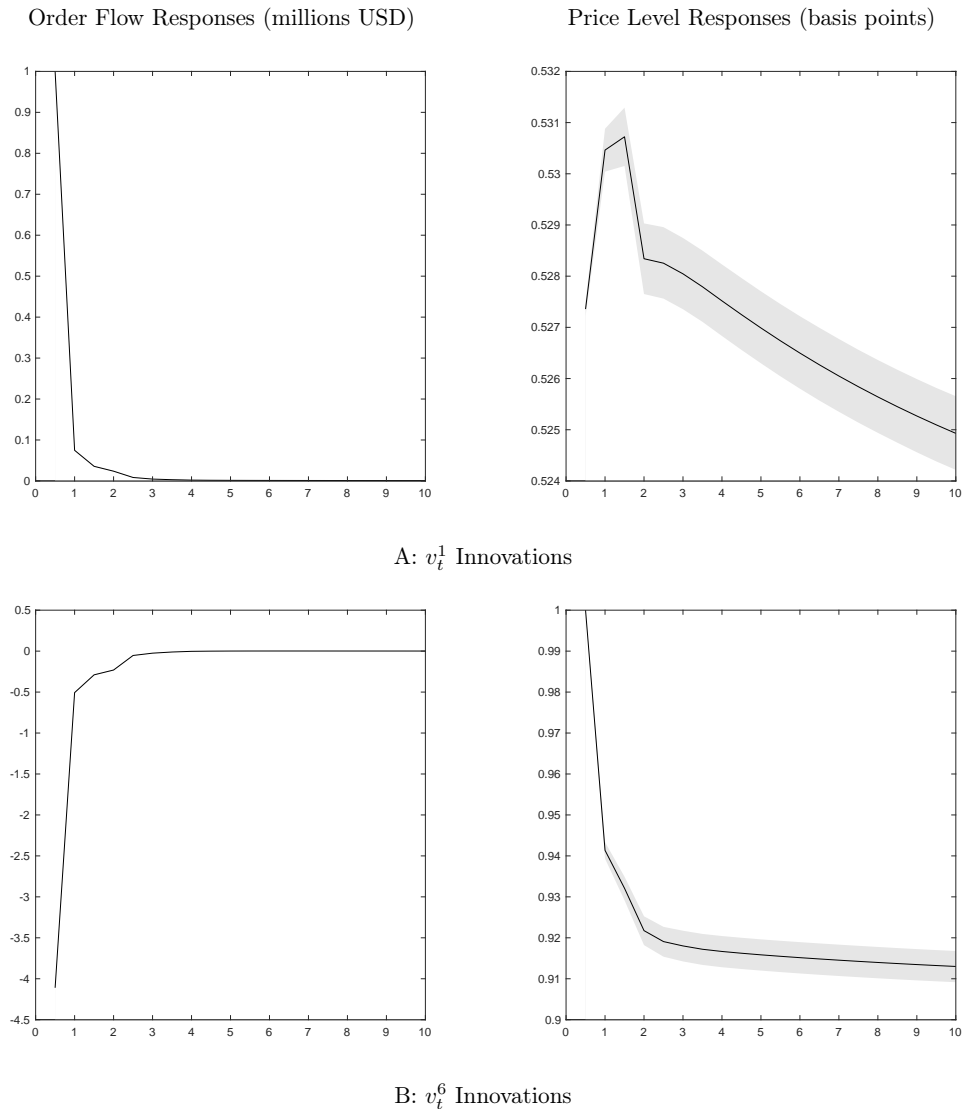
The plots in panel A of Figure 3 show the effects of a  $v_t^1$  innovation to order flow. The immediate effect of the innovation is to increase order flow by 1 million USD (by construction). Thereafter, the effects of the innovation on order flow quickly diminish, becoming negligible after three minutes. The right-hand plot in panel A shows the response of the price level  $p_t$ . (These responses are computed by cumulating the price change responses from the VAR.) The immediate price-impact is to increase the price level by 0.53 basis points. Prices rise further for the next minute and then begin a gradual decline. This decline continues beyond the ten minutes shown in the plot. According to the estimates of  $\Psi$  matrix, the impulse response reaches a value of 0.521. Thus the model estimates imply that the permanent price-impact of the 1 million USD  $v_t^1$  innovation is 0.52 basis points. This response pattern is qualitatively similar to the response of the price level to exogenous order flow shocks in the trading model shown in Figure 1. This suggests that order flow shocks affect prices via *both* the information and liquidity channels; and that the latter channel has non-negligible effects for many minutes following the shock.

The plots in panel B show the responses to a  $v_t^6$  innovation that immediately increases the price level by one basis point. It is immediately clear from these plots that news shocks cannot be the primary contributor to such innovations. Theoretically speaking, news should not affect order flow, and its price-impact should be

<sup>13</sup>The Appendix contains a complete set of impulse responses.

immediate and complete. In contrast, Figure 3 shows that the  $v_t^6$  innovation induces an immediate negative order flow of approximately four million USD, which then dies out in the following two minutes. During this period, the price level falls by approximately 0.08 basis points. Thereafter, the price response approaches its long-run value of 0.91 more slowly. The initial effects of the  $v_t^6$  innovation are qualitatively similar to the impact of the balance shock: the innovation induces a negative correlation between unexpected prices and order flow. In the minutes that follow, the effects of the  $v_t^6$  innovation on prices and order flow die out more quickly than in the trading model.

**Figure 3: Impulse Responses**



Notes: The plots show the impulse response of order flow and the price level to a  $v_t^1$  innovation (panel A) and  $v_t^6$  innovation (panel B). The shaded bands around each response are 95% confidence bands computed from a bootstrap. The horizontal axis shows the number of minutes since the innovations.

Overall, the impulse responses in Figure 3 provide empirical support for two key predictions of the trading model. First, they show that contemporaneous co-movements between order flow and price changes vary according to the shock affecting trading behavior. Second, it is important to distinguish between the immediate price-impact of a shock and its permanent price impact. These findings have important implications for understanding the source of FX price movements and order flows.

Table 3 reports the estimated variance contributions of the  $v_t^1$  and  $v_t^6$  innovations to order flow and price changes over horizons ranging from one minute to one hour, together with the lower and upper bounds of the 95% confidence intervals for each estimated contribution computed from the bootstrap. Panel A shows that approximately 61% of the variance in order flow is attributable to  $v_t^1$  innovations and approximately 39% to  $v_t^6$  innovations across all horizons. From the perspective of the trading model, these estimates suggest that variations in order flow are not simply due to exogenous shocks to the flows of market orders, but they also reflected the arrival of new information about the state of the LOB. The statistics in panel B show that  $v_t^1$  innovations are the dominant driver of FX price changes, accounting for approximately 87% of the variance across all horizons. In contrast, the variance contribution of  $v_t^6$  innovations is relatively unimportant at approximately 12%. (Together, the  $v_t^1$  and  $v_t^6$  innovations account for 99% of the variance in FX price changes.) It is also worth noting that 95% confidence intervals around all the estimated variance contributions in the table are very tight.

**Table 3: Variance Decomposition**

Variable	Innovation					
	$v_t^1$ (Order Flow)			$v_t^6$ (Balance)		
	lower	estimate	upper	lower	estimate	upper
A: Order Flow						
horizon						
1	(60.830%	60.871%	60.913%)	(39.087%	39.129%	39.170%)
2	(60.511%	60.553%	60.596%)	(39.254%	39.296%	39.338%)
10	(60.334%	60.376%	60.418%)	(39.383%	39.426%	39.468%)
60	(60.311%	60.353%	60.396%)	(39.367%	39.410%	39.453%)
B: Price Change						
horizon						
1	(87.050%	87.057%	87.065%)	(11.905%	11.911%	11.916%)
2	(86.982%	86.990%	86.998%)	(11.936%	11.942%	11.948%)
10	(86.971%	86.979%	86.987%)	(11.936%	11.943%	11.949%)
60	(86.964%	86.972%	86.981%)	(11.936%	11.942%	11.948%)

Notes: The columns headed estimate report the percentage contributions of  $v_t^1$  and  $v_t^6$  innovations to order flow and price changes over horizons ranging from 1 to 60 minutes. The columns headed lower and upper report the lower and upper boundaries of the 95% confidence interval for each variance contribution computed from a bootstrap. Estimated variance contributions and bootstrap boundaries are based on estimates of a third-order VAR estimated at a 30-second frequency, over 13 years; January 6, 2003 to December 31, 2015.

The results in the Table 3 are quite surprising. To understand why it is worth considering what we would expect to see if conventional views about FX price dynamics were true. Consider, first, the textbook

macroeconomic view that FX prices primarily react to news concerning future exchange rate fundamentals, such as interest rates or their macroeconomic determinants (see, e.g., Engel and West 2006, Engel et al., 2008 and Mark, 2009). In this case, the  $v_t^6$  innovations would primarily reflect such news, so we would expect them to contribute little to the variance of order flow and almost 100% to the variance of FX price changes at all horizons. Obviously, the results in Table 3 are very different.

An alternative view explains the contemporaneous short-horizon correlation between order flow and price changes while retaining the focus on fundamentals' news as the most important driver of FX prices over long horizons (see, e.g., Froot and Ramadorai, 2005, Berger et al., 2008a and Breedon and Vitale, 2010). According to this view, shocks to order flow only have a price impact via the liquidity channel. The operation of this channel accounts for the contemporaneous correlation between order flow and price changes over short horizons. It also means that shocks to order flow have no permanent price impact. In this case, the  $v_t^1$  innovations should account for all of the order flow variance across horizons, and some of the variance of price changes at short horizons. At long horizons, the  $v_t^6$  innovations (representing news) should account for almost 100% of the variance of FX price changes (as in the textbook macroeconomic view). Again, the results in table 3 are very different.

This is not to say that liquidity effects are absent. But, if the liquidity channel were important, we should expect to see that  $v_t^1$  innovations make a larger contribution to the variance of price changes at short horizons than at long horizons. The estimated variance contributions in Table 3 display this pattern, but the difference between the contributions at the short and long horizons are very small. This suggests that while the liquidity channel is operative, its quantitative impact on the short-term price impact is very small.

Table 3 contains two further noteworthy results. The first concerns the size of the variance contributions of  $v_t^6$  innovations to order flow. From the perspective of the trading model,  $v_t^6$  innovations represent revisions in the estimated state of the LOB so they contribute to the variance of order flow insofar as traders submitting market orders have an informational advantage over traders with existing limit orders. The estimated variance contributions of approximately 39% indicate that this endogenous response of order flow is empirically important. The second noteworthy result concerns the estimated 87% contribution of  $v_t^1$  innovations to the variance of FX price changes. Earlier research by Payne (2003), Danielsson and Love (2006), Berger et al. (2008b) and others found that order flows accounted for no more than 50% of the variance of intraday changes in FX prices. One way to account for this difference is to view order flow as a noisy signal of the "informative" flow that carries the information driving intraday FX price changes.

### 3 Macro Empirical Analysis

The results from the VAR show that intraday movements in FX prices (i.e., the EURUSD rate) are overwhelmingly driven by information that is conveyed by order flows, but not all order flows are equally important in conveying information. In this section, I use the VAR as a filter to identify the different components driving daily and weekly movements in EURUSD rates. I then examine how these components are linked to the behavior of interest rates.

### 3.1 VAR Filtering

I use the VAR estimates to decompose the intraday FX price changes and order flows into two components: one driven by the  $v_t^1$  innovations, and ones driven the  $v_t^6$  innovations. I then aggregate the price change and order flow components into daily and weekly time series.

To begin, I compute the implied vector moving average process for  $Y_{t+1}$  from the VAR estimates as

$$Y_{t+1} = \hat{\Theta}(L)\hat{V}_{t+1}, \quad (21)$$

where  $\hat{V}_{t+1} = [\hat{v}_{t+1}^1, \dots, \hat{v}_{t+1}^6]$  is the vector of estimated innovations, and  $\hat{\Theta}(L) = \hat{\Theta}_0 + \hat{\Theta}_1 L + \hat{\Theta}_2 L^2 + \dots$  is the estimated moving average polynomial with estimated matrices  $\hat{\Theta}_i$ . The contribution of the  $v_t^j$  innovation to the vector  $Y_{t+1}$  can now be calculated for every 30-second period as  $Y_{t+1}^j = \hat{\Theta}(L)\hat{V}_{t+1}^j$ , where the vector  $\hat{V}_{t+1}^j = [0, \dots, \hat{v}_{t+1}^j, \dots, 0]'$  only contains the estimated  $v_t^j$  innovation. Our interest is in the contributions of the  $v_t^1$  and  $v_t^6$  innovations to order flows and prices. I compute the contributions to order flow as  $flw_t^1 = [1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0] Y_{t+1}^1$  and  $flw_t^6 = [1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0] Y_{t+1}^6$ , and the contributions to price changes as  $\Delta p_{t+1}^1 = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1] Y_{t+1}^1$  and  $\Delta p_{t+1}^6 = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1] Y_{t+1}^6$ . To construct daily time series from these components, I choose 17:00 hrs as the end of a day. The daily order flow component driven by  $v_t^j$  innovations is just the sum of the 30-second order flow components in the 24 hours before 17:00 hrs on day  $T$ :  $flw_T^j = \sum_{t \in \text{day}(T)} flw_t^j$ . Similarly, I compute the daily price change (or depreciation rate) component driven by  $v_t^j$  innovations as  $\Delta p_T^j = \sum_{t \in \text{day}(T)} \Delta p_t^j$ . Hereafter, I use “ $T$ ” subscripts to index daily time series.

Table 4 reports summary statistics for daily EURUSD order flow and its components, and the daily EURUSD depreciation rate and its components. Panel I shows that the order flow component driven by the  $v_t^6$  innovations is more variable than the component driven by the  $v_t^1$  innovations and that the two components are positively correlated at the daily level. This pattern is repeated across the three sub-periods, but the  $flw_T^6$  component is much more volatile than the  $flw_T^1$  component during the 2007-2011 period. Panel II shows a somewhat different pattern. Here the depreciation components are negatively correlated and the  $\Delta p_T^1$  component is more variable than the  $\Delta p_T^6$  component.

The correlations between the daily components deserves comment. By construction, there is no contemporaneous correlation between the intraday components of order flow and price changes: i.e.,  $Corr(flw_t^1, flw_t^6) = 0$  and  $Corr(\Delta p_t^1, \Delta p_t^6) = 0$ . The estimated intraday innovations  $\hat{v}_t^1$  and  $\hat{v}_t^6$  are also uncorrelated with the three lagged values of  $Y_t$  that appear in the estimated VAR. This means that order flow or price change components aggregated over a couple of minutes are uncorrelated by construction. This reasoning does not apply to the daily order flow and depreciation components because estimated innovations from later in the day can be correlated with price changes and order flows that were observed hours earlier. This could occur, for example, if traders decided to submit market or limit orders based on their analysis of order flow and price movements in the previous hour. Obviously, these correlations would not exist if the estimated innovations were computed from a VAR with hundreds of lags (that spanned at least 24 hours), but estimating such VAR is completely impractical. Thus, the correlations between the daily components reported in Table 4 are indicative of inter-temporal links between price changes and order flow within the day that are undetectable

**Table 4: Summary Statistics for EURUSD Daily Series**

Sample	I: Order Flows				II: Depreciation Rates			
	$Std(flw_T)$	$Std(flw_T^1)$	$Std(flw_T^6)$	$Corr$	$Std(\Delta p_T)$	$Std(\Delta p_T^1)$	$Std(\Delta p_T^6)$	$Corr$
A: 2003-2015 (3388 obs)	394.090	142.490	308.310	0.448	63.117	68.896	39.060	-0.449
B: 2003-2006 (1039 obs)	233.980	114.570	179.020	0.118	59.451	55.443	22.678	-0.119
C: 2007-2011 (1305 obs)	563.700	175.300	445.690	0.585	69.595	84.746	56.466	-0.586
D: 2012-2015 (1044 obs)	226.140	119.570	165.520	0.203	57.906	57.810	20.964	-0.204

Notes: The table reports the standard deviations of daily order flows and their components (panel I) and the daily depreciation rates and their components (panel II) together with the correlation between the components (in the columns labelled “*Corr*”). The statistics in Panels A - D are computed for the whole sample period and three sub-periods.

in the intraday data.

### 3.2 Order Flows and the Transmission of Macro Information

I now examine whether the daily order flows and depreciation rates,  $flw_T^j$  and  $\Delta p_T^j$ , are linked to the flows of information concerning exchange-rate fundamentals. For this purpose, it is useful to refer to an identity that links daily changes in FX prices with expectations. The identity is derived from the definition of the expected log excess return on holding euros between the days  $T$  and  $T + 1$ :

$$er_T = E_T p_{T+1} - p_T + i_T^* - i_T, \quad (22)$$

where  $i_T$  and  $i_T^*$  are the log US and Euro one day nominal interest rates.  $E_T$  denotes expectations conditioned on information known at 17:00 hrs on day  $T$ . Re-arranging (22) gives the following expression for daily depreciation rate

$$\begin{aligned} \Delta p_{T+1} = & (i_T - i_T^* + er_T) - (E_{T+1} - E_T) \sum_{i=1}^H (i_{T+i} - i_{T+i}^*) \\ & - (E_{T+1} - E_T) \sum_{i=1}^H er_{T+i} + (\bar{p}_{T+1} - E_T \bar{p}_{T+1}), \end{aligned} \quad (23)$$

where  $\bar{p}_T \equiv E_T p_{T+H}$  is the expected EURUSD rate many (i.e.,  $H > 0$ ) days ahead.<sup>14</sup>

<sup>14</sup>The derivation of this equation requires several steps, which are described in the Appendix. Importantly, the equation follows from the Law of Iterated Expectations and the definition in (22), it embeds no additional assumptions.



Equation (6) identifies all the proximate factors that can drive the daily depreciation rate. The first term on the right identifies the expected depreciation rate,  $E_T \Delta p_{T+1}$ . The remaining terms identify the factors that contribute to the error in forecasting the EURUSD rate one day ahead, i.e.,  $p_{T+1} - E_T p_{T+1}$ . These factors are: (i) news about future interest differentials (over the next  $H$  days), (ii) news about expected excess returns (over the next  $H$  days), and (iii) revisions in expectations concerning the EURUSD rate  $H$  days ahead.

Recall that the  $v_t^1$  innovations appear to be the dominant driver of intraday changes in EURUSD rates, accounting for approximately 87% of their variance. Since daily depreciation rates have very little predictability, equation (6) implies that  $v_t^1$  innovations primarily convey news about future interest differentials, expected excess returns, and/or long-horizon FX prices. It is therefore natural to empirically investigate whether the daily order flow component driven by  $v_t^1$  innovations,  $flw_T^1$ , is related to other variables that embed news about interest differentials. The variable I use for this purpose is the daily change in the difference between US and EUR one-year nominal interest rates:  $\Delta(i_{T+1}^{(yr)} - i_{T+1}^{*(yr)})$ . Changes in these one-year rates reflect revisions in the expected path of daily rates over the next year, and changes in the risk premium embedded in the term structure of interest rates. So if we set the horizon  $H$  in equation (6) to one year,  $\Delta(i_{T+1}^{(yr)} - i_{T+1}^{*(yr)})$  should be correlated with  $(E_{T+1} - E_T) \sum_{i=1}^H (i_{T+i} - i_{T+i}^*)$ .

Table 5 shows the results of regressing the daily order flow components,  $flw_T^1$  and  $flw_T^6$ , on the daily change in the difference between the US and EUR one-year nominal interest rates,  $\Delta(i_{T+1}^{(yr)} - i_{T+1}^{*(yr)})$ . The regressions also include a constant and one lag of the dependent variable (in some cases). The table reports estimates computed from the full sample period, and from three sub-periods. These sub-periods were chosen to give results based on data before, during, and after the 2008 financial crisis.

As the table shows, there is a striking difference between the estimated coefficients in the regressions for the  $flw_T^1$  component in panel I and the estimated coefficients in the regressions for the  $flw_T^6$  in panel II. The estimated coefficient on the change in the interest differential  $\Delta(i_{T+1}^{(yr)} - i_{T+1}^{*(yr)})$  is negative and highly statistically significant in the  $flw_T^1$  regressions in both the full sample period and all but the first sub-period. In contrast, none of the estimated coefficients on  $\Delta(i_{T+1}^{(yr)} - i_{T+1}^{*(yr)})$  are statistically significant in the  $flw_T^6$  regressions. These results suggest that news concerning future interest differentials contributed to the  $v_t^1$  innovations that drove that  $flw_T^1$  component of order flow. Moreover, it appears that the order flow component fell on days where news produced an upward revisions in expectations  $(E_{T+1} - E_T) \sum_{i=1}^H (i_{T+i} - i_{T+i}^*)$ . This pattern is consistent with equation (6) because  $v_t^1$  innovations produce a positive correlation between order flow and depreciation rates (FX price changes).

To provide further evidence on the information conveyed by order flow, Table 6 reports estimates of regressions for the daily depreciation rate  $\Delta p_{T+1}$ . Panel I shows the results from regressing  $\Delta p_{T+1}$  on the change in the interest differential  $\Delta(i_{T+1}^{(yr)} - i_{T+1}^{*(yr)})$ . Some specifications also include two lags of the depreciation rate, order flow and the interest differential. As the table shows, the estimated coefficients on the interest differential are negative and statistically significant in both the full sample-period and the sub-periods. However, these regressions account for very little of the variability in the depreciation rate – the largest adjusted  $R^2$  statistics equal 0.1 in the 2007-11 sub-period, and approximately 0.05 over the

whole 13 year sample period. Panel II reports estimates of regressions that replace the change in the interest differential with the component of the depreciation rate driven by  $v_t^1$  innovations,  $\Delta p_{T+1}^1$ . In contrast to the results in panel I, the estimated coefficients on the depreciation component are positive and highly significant. The large adjusted  $R^2$  statistics confirm that  $v_t^1$  innovations are the dominant drivers of daily EURUSD depreciation rates, consistent with the variance decomposition results in Table 3.

**Table 5: Daily Order Flow Regressions**

Sample	Variable	Specification			
		I: $flow_t^1$		II: $flow_t^6$	
A: 2003-2015					
(3388 obs)	$\Delta(i^{(yr)} - i^{*(yr)})$	-3.995*** (0.613)	-3.964*** (0.602)	0.276 (0.673)	0.356 (0.660)
	lag	no	yes	no	yes
	$\bar{R}^2$	0.038	0.041	0.000	0.072
B: 2003-2006					
(1039 obs)	$\Delta(i^{(yr)} - i^{*(yr)})$	-1.044* (0.552)	-1.038* (0.559)	-0.259 (0.930)	-0.150 (0.790)
	lag	no	yes	no	yes
	$\bar{R}^2$	0.003	0.003	0.001	0.052
C: 2007-2011					
(1305 obs)	$\Delta(i^{(yr)} - i^{*(yr)})$	-4.916*** (0.651)	-4.811*** (0.641)	0.563 (0.946)	0.648 (0.958)
	lag	no	yes	no	yes
	$\bar{R}^2$	0.062	0.076	0.001	0.077
D: 2012-2015					
(1044 obs)	$\Delta(i^{(yr)} - i^{*(yr)})$	-14.677*** (1.732)	-14.677*** (1.734)	-0.919 (2.095)	-0.976 (2.070)
	lag	no	yes	no	yes
	$\bar{R}^2$	0.087	0.087	0.001	0.037

Notes: The table reports OLS estimates for regressions of the daily order flow components:  $flow_T^1$  (panel I) and  $flow_T^6$  (panel II). All regressions include a constant and daily change in the 12-month interest differential  $\Delta(i^{(yr)} - i^{*(yr)})$  as right-hand side variables. Regressions with “lag” also include the one lag of the dependent variable. Heteroskedastic-consistent standard errors are reported in parenthesis below each of the parameter estimates. Statistical significance at the 10, 5 and 1% levels is indicated by \* \*\*, and \*\*\*, respectively.

**Table 6: Daily Depreciation Rate Regressions**

Sample	Variable	Specification					
		I		II		III	
2003-2015	$\Delta(i^{(yr)} - i^{*(yr)})$	-1.979*** (0.300)	-2.028*** (0.285)				
	$\Delta p^1$			0.742*** (0.017)	0.759*** (0.017)	1.027*** (0.052)	1.029*** (0.052)
	lags	no	yes	no	yes	no	yes
	$\bar{R}^2$	0.048	0.051	0.656	0.673	0.559	0.588
2003-2006	$\Delta(i^{(yr)} - i^{*(yr)})$	-0.428*** (0.190)	-0.792*** (0.253)				
	$\Delta p^1$			0.961** (0.017)	0.971** (0.016)	0.890** (0.336)	0.987** (0.225)
	lags	no	yes	no	yes	no	yes
	$\bar{R}^2$	0.003	0.008	0.807	0.813	0.802	0.813
2007-2011	$\Delta(i^{(yr)} - i^{*(yr)})$	-2.480*** (0.329)	-2.466*** (0.323)				
	$\Delta p^1$			0.598*** (0.025)	0.619*** (0.025)	1.046*** (0.053)	1.033*** (0.056)
	lags	no	yes	no	yes	no	yes
	$\bar{R}^2$	0.101	0.102	0.530	0.552	0.232	0.306
2012-2015	$\Delta(i^{(yr)} - i^{*(yr)})$	-7.069*** (0.841)	-7.083*** (0.866)				
	$\Delta p^1$			0.929*** (0.025)	0.933*** (0.025)	1.000*** (0.037)	1.002*** (0.041)
	lags	no	yes	no	yes	no	yes
	$\bar{R}^2$	0.087	0.084	0.860	0.865	0.855	0.860

Notes: The table reports estimates of the daily depreciation rate for the EURUSD on the daily change in the 12-month interest differential  $\Delta(i - i^*)$ , and the order flow component of the daily depreciation rate  $\Delta p^{\text{flw}}$ . Regressions with “lags” also include two lags of  $\Delta(i - i^*)$ , the daily depreciation rate and daily order flow. Specifications I and II are estimated by OLS. Specification III is estimated by GMM using  $\Delta(i - i^*)$  as an instrument for  $\Delta p^{\text{flw}}$ . Heteroskedastic-consistent standard errors are reported in parenthesis below each of the parameter estimates. Statistical significance at the 10, 5 and 1% levels is indicated by \* \*\*, and \*\*\*, respectively.

The estimates in Panel III of Table 6 speak to the question of what drives the  $v_t^1$  innovations. Here I report the estimates from an instrumental variable (IV) regression of  $\Delta p_{T+1}$  on  $\Delta p_{T+1}^1$  where change in the interest differential are used as the instrument for  $\Delta p_{T+1}^1$ . If  $v_t^1$  innovations are unrelated to revisions in expectations concerning future interest rates, the IV estimates of the  $\Delta p_{T+1}^1$  coefficient should be close

to zero and insignificant. Alternatively, if changing interest rate expectations account for most of the  $v_t^1$  innovations, the IV estimates of the  $\Delta p_{T+1}^1$  coefficient should be significant and close to one. The results in the table confirm this prediction. In all cases the estimated coefficients are highly significant and close to one. Indeed, despite the small standard errors, t-tests for the null of a coefficient equal to one cannot be rejected at conventional significance levels. The adjusted  $R^2$  statistics from these regressions are also noteworthy. Across the full sample period, and two of the sub-periods, these statistics are very similar in magnitude to those in panel II. This implies that most of the variations in  $\Delta p_{T+1}^1$  during these periods were correlated with changes in the interest differential. The one exception to this pattern appears in the 2007-2011 sub-period, where the adjusted  $R^2$  statistics in panel III are a good deal smaller than those in panel II. It appears that  $v_t^1$  innovations conveyed less information about revisions in interest-rate expectations and were a less dominant source of daily FX price movements during this sub-period. Perhaps this is not too surprising in light of the fact that much of the 2007-2011 period includes the 2008 financial crisis.

The results in Tables 5 and 6 are notable from several perspectives. First, they contradict the view that spot exchange rates are largely disconnected from macro fundamentals; a view Obstfeld and Rogoff (2000) called the Exchange-Rate Disconnect Puzzle. If we exclude the sub-period dominated by the 2008 financial crisis, at least 80% of the daily variations in the EURUSD spot rate are attributable to order flows correlated with changing long-term interest rates. It appears that order flow, or more precisely a component of order flow, acts as the medium connecting spot rates with changing expectations about future short-term interest rates.<sup>15</sup> Second, the results appear to confirm earlier research linking order flows to macro information. Evans (2010) and Evans and Lyons (2013) found that order flows from end-users contain incremental (non-public) information about (unreported) current and future macro variables, while Rime et al. (2010) show that order flows have short-term forecasting power for specific macro data releases. The results here use a component of order flow from the wholesale market rather than the end-user flows of an individual bank, and cover a much longer time period.

Finally, it is worth considering why the regressions with order flow in panels II and III of Table 6 are so much more successful in accounting for daily depreciation rates than the regressions in panel I that only include changes in the interest differential,  $\Delta(i_{T+1}^{(yr)} - i_{T+1}^{*(yr)})$ . Suppose changes in the interest differential were entirely due to revisions in expectations about future short-term interest rates, and those revisions were driven by the arrival of public news  $n_t$  as identified in the trading model. Under these circumstances, the regressions in panel I of Table 6 would have high adjusted  $R^2$  statistics. Furthermore, the regressions of the daily order flow components on  $\Delta(i_{T+1}^{(yr)} - i_{T+1}^{*(yr)})$  in Table 5 would produce insignificant coefficients. The fact that neither of these predictions is borne out empirically suggests that revisions in expectations are not primarily driven by such news. It seems, instead, that throughout the day expectations are revised through an unobserved process that also produces (what appear as) shocks to order flow, which ultimately are transmitted to FX prices through the information channel. Thus the  $flw_T^1$  component of order flow

---

<sup>15</sup>Of course, movements in long-term interest rates could reflect changes in the term premia rather than expectations concerning future short-term rates, so it is *possible* that the  $v_t^1$  innovations conveyed information about term premia. I regard this as a less likely explanation for the results in panel III because the explanatory power of order flow is so much less between 2007 and 2011 than in the earlier or later sub-periods.

appears as the key link connecting movements in spot exchange rates to macro fundamentals via changing expectations concerning future interest rates.

### 3.3 Order Flows and Risk Premia

If the  $flw_T^1$  component of order flow carries information about future interest rates, what is the role of the  $flw_T^6$  component? To address this question, let us return to the identity in equation (23), which I now rewrite as

$$\begin{aligned} & \Delta p_{T+1} + (E_{T+1} - E_T) \sum_{i=1}^H (i_{T+i} - i_{T+i}^*) \\ &= E_T \Delta p_{T+1} - (E_{T+1} - E_T) \sum_{i=1}^H er_{T+i} + (\bar{p}_{T+1} - E_T \bar{p}_{T+1}). \end{aligned} \quad (24)$$

The results in Panel III of Table 6 imply that  $\Delta p_{T+1}^1$  is a good proxy for  $-(E_{T+1} - E_T) \sum_{i=1}^H (i_{T+i} - i_{T+i}^*)$ , so the term on the left-hand-side of (24) is approximately equal to  $\Delta p_{T+1} - \Delta p_{T+1}^1 = \Delta p_{T+1}^6$ . Thus the  $flw_T^6$  component must reflect the expected depreciation rate  $E_T \Delta p_{T+1}$ , revisions in expected future excess returns  $(E_{T+1} - E_T) \sum_{i=1}^H er_{T+i}$ , and/or changing long-horizon forecasts for the EURUSD rate  $\bar{p}_{T+1} - E_T \bar{p}_{T+1}$ .

To differentiate between these alternative possibilities, I estimate forecasting regressions of the form

$$\Delta^\tau p_{T+\tau}^6 = \alpha_0 + \alpha_1 Z_{1,T} + \dots + \alpha_j Z_{j,T} + u_{T+\tau}, \quad (25)$$

where  $\Delta^\tau p_{T+\tau}^6 = \sum_{j=1}^T \Delta p_{T+j}^6$  is the  $\tau$ -day change in  $p_T^6$  and  $Z_{j,T}$  are forecasting variables known on day  $T$ . If movements in the  $flw_T^6$  component reflect changing expectations about depreciation rates, and those expectations are correlated with the forecasting variables, some of the  $\alpha_j$  coefficients should be statistically significant.

I consider three forecasting variables: the change in the interest differential on one week bonds  $\Delta(i_T^{(w)} - i_T^{*(w)})$ , the  $\Delta p_T^6$  component of the daily depreciation rate, and the realized skewness in intraday FX price changes during day  $T$ :  $skw_T = \sum_{i \in \text{day}(T)} (\Delta p_i)^3$ . The trading model from Section 1 motivates these forecasting variable choices. Recall that the impulse responses to the  $v_t^6$  innovations in Figure 3 resembled the responses in the trading model to shocks that change the estimated depth balance. If  $v_t^6$  innovations are indeed primarily driven by revisions in  $E_t bal_t$ , the  $\Delta p_T^6$  component of the daily depreciation should have forecasting power for  $\Delta^\tau p_{T+\tau}^6$ .

To understand the economic logic behind this prediction, we need to consider the risks associated with the submission of a limit order. A trader submitting a limit order to sell FX faces two risks. The first is the risk that the order will not execute (i.e., it will not be matched with an incoming market buy order). The second risk is that the order will execute along with other limits with much higher offer prices (because there is a large flow of market buy orders). In this case the trader would have missed the opportunity to sell FX at a higher price. The size of both risks depends on the probability distribution for the flow of market buy orders used by traders. Suppose this distribution shifts to the right, so the risk of underpricing the limit sell

order rises, and the risk of non-execution falls. *Ceteris paribus*, such a shift will reduce the depth of limit sell orders, which in turn will raise the actual and estimated depth balance. Analogous reasoning applies to limit buy orders. In this case a rightward shift in the distribution for the flow market sell orders would lower the depth of limit buy orders and lower  $E_t bal_t$ , *ceteris paribus*. When shifts in the distribution for market buy and sell orders reflect changes in the distribution of order flow shocks, movements in  $E_t bal_t$  will reflect changes in the distribution of future FX prices. In particular, if traders believe that the distribution of order flow shocks has shifted to the right,  $E_t bal_t$  will rise and the distribution of future FX price changes will shift to the right (because large positive order flow shocks are more likely). As the trading model showed, an upward revision in  $E_t bal_t$  produces an increase in FX prices. So  $\Delta p_T^6$  should be positive on days where the consensus view among traders is that distribution of future FX price changes has moved to the right. Insofar as this consensus view is correct, positive (negative) values for  $\Delta p_T^6$  should forecast higher (lower) future changes in FX prices.

The results from estimating the forecasting regression in (25) are reported in Table 7. Panel I shows the results from forecasting the daily depreciation component  $\Delta p_{T+1}^6$  one day ahead, while Panel II shows the results from forecasting the weekly depreciation component  $\Delta^5 p_{T+5}^6$  one week ahead. As above, I report estimates for the whole sample and three sub-periods. Overall, the table shows that the  $\Delta p_T^6$  component has forecast power at both horizons. The estimated coefficients on  $\Delta p_T^6$  are positive and highly statistically significant in all the specifications. In contrast, changes in the interest differentials appear to have no forecasting power. This finding is robust to the inclusion of the interest differential rather than the change in the interest differential. The table also shows that the forecasting power of  $\Delta p_T^6$  is not diminished by the addition of realized skewness  $skw_T$ . If traders changed their views about the future distribution of price changes based on past changes, the forecasting power of  $\Delta p_T^6$  could disappear when this variable is added to the regression. As the table shows, the  $skw_T$  variable appears significant in some of the regressions, but the forecasting power of  $\Delta p_T^6$  appears robust to its inclusion.

The forecasting results presented here distinct from those found in earlier studies. For example, Evans and Lyons (2005) and Evans and Lyons (2013) use the end-user order flows received by Citibank to forecast EURUSD depreciation rates (both within-sample and out-of-sample) up to four weeks ahead.<sup>16</sup> These order flows represent private information that was not available across the market, whereas the order flow components used here come from trades on the EBS LOB that are observable to a large number of market participants.

Of course there is a long tradition in exchange-rate economics, starting with Meese and Rogoff (1983), of considering the forecasting implication of different models. The results here add two perspectives to this tradition. First, because a component of order flow appears to convey new information about future interest rates that typically accounts for a large fraction of the variance in daily depreciation rates, the ability of any variable (known *ex ante*) to forecast daily depreciation rates should be quite limited. Second, the theoretical foundation of the forecasting results presented here considers the risks associated with the provision of

---

<sup>16</sup>In contrast to Evans and Lyons results, Reitz et al. (2011) were unable to find forecasting power in the end-user flows they studied. It appears that not all end-user order flows are equally informative.

liquidity in FX trading rather than the risk identified by the behavior of pricing kernels (stochastic discount factors) found in traditional asset-pricing models (see, e.g., Engel, 2013).

**Table 7: Forecasting Order Flow Regressions**

Sample	Forecasting Variables	Forecast Variables					
		I: One Day Ahead			II: One Week Ahead		
		$\Delta^1 p_{T+1}^6$			$\Delta^5 p_{T+5}^6$		
A: 2003-2015							
	$\Delta(i_T - i_T^*)$	0.587 (1.165)	1.004 (1.213)	0.899 (1.197)	-0.116 (0.520)	-0.125 (0.499)	-0.049 (0.513)
	$\Delta^\tau p_T^6$		0.270*** (0.032)	0.250*** (0.032)		0.207*** (0.052)	0.390*** (0.106)
	$skw_T$			0.074** (0.030)			-0.206** (0.102)
	$\bar{R}^2$	0.000	0.072	0.077	0.000	0.042	0.051
B: 2003-2006							
	$\Delta(i_T - i_T^*)$	2.857 (2.982)	1.899 (2.952)	1.922 (2.948)	0.329 (1.186)	0.086 (1.126)	0.130 (1.141)
	$\Delta^\tau p_T^6$		0.231*** (0.037)	0.229*** (0.038)		0.333*** (0.061)	0.278*** (0.089)
	$skw_T$			0.008 (0.034)			0.061 (0.086)
	$\bar{R}^2$	0.000	0.052	0.051	-0.001	0.109	0.109
C: 2007-2011							
	$\Delta(i_T - i_T^*)$	0.506 (1.274)	1.027 (1.319)	0.888 (1.297)	-0.221 (0.574)	-0.238 (0.536)	-0.124 (0.560)
	$\Delta^\tau p_T^6$		0.280*** (0.039)	0.258*** (0.040)		0.189*** (0.062)	0.419*** (0.128)
	$skw_T$			0.078** (0.037)			-0.258** (0.121)
	$\bar{R}^2$	-0.001	0.077	0.082	-0.001	0.034	0.047
D: 2012-2015							
	$\Delta(i_T - i_T^*)$	-0.379 (4.221)	-0.002 (4.209)	-0.025 (4.217)	4.420 (4.558)	4.945 (4.571)	5.015 (4.542)
	$\Delta^\tau p_T^6$		0.197*** (0.031)	0.180*** (0.032)		0.148*** (0.049)	0.210*** (0.068)
	$skw_T$			0.086*** (0.031)			-0.071 (0.075)
	$\bar{R}^2$	-0.001	0.037	0.043	0.002	0.022	0.022

Notes: The table reports estimates of forecasting regression (25). Regressions are estimated by OLS at the daily frequency. Heteroskedastic-consistent standard errors are reported in parenthesis below each of the parameter estimates. These standard errors also account for the MA(4) process induced by the overlapping forecast errors in the Panel II regressions. Statistical significance at the 10, 5 and 1% is signified by \*, \*\* and \*\*\*, respectively.

## 4 Conclusion

This paper looked inside the “black box” of FX trading in an attempt to reconnect short-term movements in spot exchange rates with macroeconomic variables. Using insights from the microstructure model of FX trading on a LOB, I used a structural VAR to identify two distinct components of order flow in wholesale FX trading that play distinctive roles in driving intraday changes in the EURUSD rate. The  $v_t^1$  innovations driving one component are the dominant driver of EURUSD rates, accounting for up to 80% of the variations in daily depreciation rates before 2007 and after 2011. Furthermore, both the daily order flow and depreciation components driven by the  $v_t^1$  innovations are strongly correlated with daily changes in long-term interest differentials, suggesting that these innovations carry information about future short-term interest rates. These findings suggest that short-term movements in EURUSD rates are often much better connected to macro variables than is widely thought.

## References

- Berger, D., Chaboud, A., Chernenko, S., Howorka, E., Wright, J., 2008a. Order flow and exchange rate dynamics in electronic brokerage system data. *Journal of International Economics* 75 (1), 93–109.
- Berger, D. W., Chaboud, A. P., Chernenko, S. V., Howorka, E., Wright, J. H., 2008b. Order flow and exchange rate dynamics in Electronic Brokerage System data 75 (1), 93–109.
- Bjønnes, G. H., Osler, C. L., Rime, D., 2007. Asymmetric information in the foreign exchange market. typescript, Norges Bank.
- Bjønnes, G. H., Rime, D., Solheim, H. O. A., Mar. 2005. Liquidity provision in the overnight foreign exchange market. *Journal of International Money & Finance* 24 (2), 177–198.
- Brandt, M. W., Kavajecz, K. A., 2004. Price discovery in the us treasury market: The impact of orderflow and liquidity on the yield curve. *The Journal of Finance* 59 (6), 2623–2654.
- Breedon, F., Vitale, P., 2010. An empirical study of portfolio-balance and information effects of order flow on exchange rates. *Journal of International Money and Finance* 29 (3), 504 – 524.
- Chinn, M., Moore, M., 2009. Private Information and the Monetary Model of Exchange Rates: Evidence from a Novel Data Set. working paper, Queens University.
- Daniélsson, J., Love, R., 2006. Feedback trading 11 (1), 35–53.
- Engel, C., 2013. Exchange rates and interest parity. Vol. 4 of *Handbook of International Economics*. Elsevier.
- Engel, C., 2014. Exchange rates and interest parity. In: *Handbook of international economics*. Vol. 4. Elsevier, pp. 453–522.



- Engel, C., Mark, N. C., West, K. D., June 2008. Exchange rate models are not as bad as you think. In: NBER Macroeconomics Annual 2007, Volume 22. University of Chicago Press, pp. 381–441.
- Engel, C., West, K., 2006. Taylor Rules and the Deutschmark–Dollar Real Exchange Rate. *Journal of Money, Credit & Banking* 38 (5), 1175–1194.
- Evans, M. D. D., 01 2010. Order flows and the exchange rate disconnect puzzle. *Journal of International Economics* 80 (1), 58–82.
- Evans, M. D. D., 2011. *Exchange-Rate Dynamics*. Princeton Series in International Finance.
- Evans, M. D. D., Lyons, R. K., Feb. 2002. Order flow and exchange rate dynamics 110 (1), 170–180.
- Evans, M. D. D., Lyons, R. K., 2005. Meese-rogooff redux: Micro-based exchange-rate forecasting. *American Economic Review. Papers & Proceedings* 95 (2), 405–414.
- Evans, M. D. D., Lyons, R. K., 2006. Understanding order flow 11 (1), 3–23.
- Evans, M. D. D., Lyons, R. K., 2013. Exchange rate fundamentals and order flow. *Quarterly Journal of Finance* 02 (04), 1250018.
- Evans, M. D. D., Rime, D., 2012. Micro approaches to foreign exchange determination. In: James, J., Marsh, I., Sarno, L. (Eds.), *Handbook of Exchange Rates*. John Wiley & Sons.
- Foucault, T., 1999. Order flow composition and trading costs in a dynamic limit order market 2, 99–134.
- Foucault, T., Pagano, M., Roell, A., Röell, A., 2013. *Market liquidity: theory, evidence, and policy*. Oxford University Press.
- Froot, K. A., Ramadorai, T., 2005. Currency returns, intrinsic value, and institutional-investor flows. *Journal of Finance* 60 (3), 1535–1566.
- Glosten, L. R., Sep. 1994. Is the electronic open limit order book inevitable 49 (4), 1127–1161.
- Gyntelberg, J., Loretan, M., Subhanij, T., Chan, E., Aug. 2009. International portfolio rebalancing and exchange rate fluctuations in Thailand. Tech. Rep. 287, BIS.
- Hasbrouck, J., 1991. Measuring the information content of stock trades. *Journal of Finance* 46 (1), 179–207.
- Hasbrouck, J., 1996. Modeling market microstructure time series. *Handbook of statistics* 14, 647–692.
- Hau, H., Killeen, W., Moore, M., Apr. 2002. How has the euro changed the foreign exchange market? 17 (34), 149–192.
- Ito, T., Yamada, M., 2015. Was the forex fixing fixed? Tech. rep., National Bureau of Economic Research.
- Killeen, W. P., Lyons, R. K., Moore, M. J., Jun. 2006. Fixed versus flexible: Lessons from EMS order flow 25 (4), 551–579.

- King, M., Osler, C., Rime, D., 2011. Foreign exchange market structure, players and evolution.
- King, M., Sarno, L., Sojli, E., Dec. 2010. Timing exchange rates using order flow: The case of the Loonie. *Journal of Banking & Finance* 34 (12), 2917–2928.
- Love, R., Payne, R., 2008. Macroeconomic news, order flows and exchange rates 43 (02), 467–488.
- Lyons, R. K., 1995. Tests of microstructural hypothesis in the foreign exchange market 39 (2-3), 321–351.
- Lyons, R. K., 1997. A simultaneous trade model of the foreign exchange hot potato 42 (3-4), 275–298.
- Lyons, R. K., 2001. *The Microstructure Approach to Exchange Rates*. Cambridge, MA.
- Mark, N., 2009. Changing Monetary Policy Rules, Learning, and Real Exchange Rate Dynamics. *Journal of Finance & Banking* 41 (6), 1047–1070.
- Marsh, I. W., O'Rourke, C., 2005. Customer order flow and exchange rate movements: Is there really information content? Working paper, Cass Business School.
- Meese, R. A., Rogoff, K., February 1983. Empirical exchange rate models of the seventies : Do they fit out of sample? *Journal of International Economics* 14 (1-2), 3–24.
- Menkhoff, L., Schmeling, M., 2008. Local information in foreign exchange markets 27 (8), 1383–1406.
- Obstfeld, M., Rogoff, K., 2000. The six major puzzles in international macroeconomics: is there a common cause? *NBER macroeconomics annual* 15, 339–390.
- O'Hara, M., 1997. *Market microstructure theory*. Wiley.
- Onur, E., 2008. The role of asymmetric information among investors in the foreign exchange market 13 (4), 368–385.
- Osler, C. L., 2009. Market microstructure, foreign exchange. In: Meyers, R. A. (Ed.), *Encyclopedia of Complexity and System Science*. Springer, New York, pp. 5404–5438.
- Osler, C. L., Mende, A., Menkhoff, L., 2011. Price discovery in currency markets. *Journal of International Money and Finance* 30 (8), 1696 – 1718.
- Parlour, C. A., Seppi, D. J., 2008. Limit order markets: A survey. *Handbook of financial intermediation and banking* 5, 63–95.
- Payne, R., Dec. 2003. Informed trade in spot foreign exchange markets: An empirical investigation 61 (2), 307–329.
- Reitz, S., Schmidt, M. A., Taylor, M. P., 2011. End-user order flow and exchange rate dynamics — a dealer's perspective. *European Journal of Finance* 17 (2), 153–168.  
 URL <http://www.tandfonline.com/doi/abs/10.1080/13518471003651925>

Rime, D., 2001. Trading in foreign exchange markets. Ph.D dissertation, Norwegian School of Management, Norway.

Rime, D., Sarno, L., Sojli, E., 2010. Exchange rate forecasting, order flow and macroeconomic information. Journal of International Economics 80 (1), 72–88.

Wu, T., 2007. Order flow in the south: Anatomy of the brazilian fx market. typescript, UC Santa Cruz.

## A Appendix

### A.1 Solution to the Trading Model

This appendix describes the solution of the microstructure trading model in equations (4) - (14). I solve the model in two steps. In the first step I find the solution for price changes, order flow and the depth balance. I then solve for for the spread, volume and total depth in the second step.

To begin, I rewrite order flow and the change in FX prices as

$$flw_{t+1} = -\alpha\beta E_t bal_t - 2\alpha\omega E_t^* \Delta p_{t+1} + u_{t+1}^{flw} \quad \text{and} \quad (26)$$

$$\begin{aligned} \Delta p_{t+1} = n_{t+1} + \lambda(flw_{t+1} - E_t flw_{t+1}) + \frac{1}{2}\beta\kappa^{bal}(bal_t + \xi_t^{bal} - E_t bal_t) \\ + \frac{1}{2}\beta flw_{t+1} - \frac{1}{2}\beta(1 - \varphi^{bal})(flw_{t+1} - E_t^* flw_{t+1}). \end{aligned} \quad (27)$$

These equations imply that  $E_t \Delta p_{t+1} = \frac{1}{2}\beta E_t flw_{t+1}$  and  $E_t flw_{t+1} = -\alpha\beta E_t bal_t - 2\alpha\omega E_t \Delta p_{t+1}$ . Combining these expressions gives

$$E_t flw_{t+1} = -\frac{\alpha\beta}{1 + \alpha\beta\omega} E_t bal_t \quad \text{and} \quad E_t \Delta p_{t+1} = -\frac{1}{2}\beta \frac{\alpha\beta}{1 + \alpha\beta\omega} E_t bal_t. \quad (28)$$

Equations (26) and (27) also imply that

$$\begin{aligned} E_t^* flw_{t+1} = -\alpha\beta E_t bal_t - 2\alpha\omega E_t^* \Delta p_{t+1} \quad \text{and} \\ E_t^* \Delta p_{t+1} = (\lambda + \frac{1}{2}\beta) E_t^* flw_{t+1} + \frac{\lambda\alpha\beta}{1 + \alpha\beta\omega} E_t bal_t + \frac{1}{2}\beta\kappa^{bal}(bal_t + \xi_t^{bal} - E_t bal_t). \end{aligned}$$

Writing these expressions in matrix form, and solving for the expectations gives

$$\begin{bmatrix} E_t^* \Delta p_{t+1} \\ E_t^* flw_{t+1} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \frac{\beta}{1 + \alpha\omega(2\lambda + \beta)} \kappa^{bal} & -\frac{1}{2}\beta \frac{\alpha\beta}{1 + \alpha\beta\omega} \\ -\alpha\omega \frac{\beta}{1 + \alpha\omega(2\lambda + \beta)} \kappa^{bal} & -\frac{\alpha\beta}{1 + \alpha\beta\omega} \end{bmatrix} \begin{bmatrix} bal_t + \xi_t^{bal} - E_t bal_t \\ E_t bal_t \end{bmatrix}. \quad (29)$$

With these results I can rewrite the order flow equation (26) as

$$flw_{t+1} = -(1 - \rho) E_t bal_t + u_{t+1}^{flw} - \omega\phi\kappa^{bal} (bal_t - E_t bal_t + \xi_t^{bal}), \quad (30)$$

where  $\rho = \frac{1}{1+\alpha\beta}$  and  $\phi = \frac{\alpha\beta}{1+\alpha\omega(2\lambda+\beta)}$ . This equation implies that  $flw_{t+1} - E_t^* flw_{t+1} = u_{t+1}^{\text{flw}}$  and  $flw_{t+1} - E_t flw_{t+1} = u_{t+1}^{\text{flw}} - \omega\phi\kappa^{\text{bal}}(bal_t - E_t bal_t + \xi_t^{\text{bal}})$ . So I can now rewrite the price change equation in (27) as

$$\Delta p_{t+1} = n_{t+1} + (\lambda + \frac{1}{2}\beta\varphi^{\text{bal}})u_{t+1}^{\text{flw}} + \frac{1}{2}[\beta - \omega\phi(\beta + 2\lambda)]\kappa^{\text{bal}}(bal_t - E_t bal_t + \xi_t^{\text{bal}}) - \frac{1}{2}\beta(1 - \rho)E_t bal_t. \quad (31)$$

To complete the first step, I solve for the estimated depth balance and its associated estimation error. Equations (5), (11) and (13) imply that

$$E_{t+1} bal_{t+1} = E_t bal_t + \kappa^{\text{bal}}(bal_t + \xi_t^{\text{bal}} - E_t bal_t) + flw_{t+1} - (1 - \varphi^{\text{bal}})(flw_{t+1} - E_t^* flw_{t+1}), \quad \text{and} \quad (32)$$

$$bal_{t+1} - E_{t+1} bal_{t+1} = (1 - \kappa^{\text{bal}})(bal_t - E_t bal_t) + u_{t+1}^{\text{bal}} - \kappa^{\text{bal}}\xi_t^{\text{bal}} + (1 - \varphi^{\text{bal}})(flw_{t+1} - E_t^* flw_{t+1}). \quad (33)$$

Simplifying these expressions gives

$$\begin{aligned} \begin{bmatrix} E_{t+1} bal_{t+1} \\ bal_{t+1} - E_{t+1} bal_{t+1} \end{bmatrix} &= \begin{bmatrix} \rho & (1 - \omega\phi)\kappa^{\text{bal}} \\ 0 & 1 - \kappa^{\text{bal}} \end{bmatrix} \begin{bmatrix} E_t bal_t \\ bal_t - E_t bal_t \end{bmatrix} \\ &+ \begin{bmatrix} \varphi^{\text{bal}} & 0 & (1 - \omega\phi)\kappa^{\text{bal}} \\ 1 - \varphi^{\text{bal}} & 1 & -\kappa^{\text{bal}} \end{bmatrix} \begin{bmatrix} u_{t+1}^{\text{flw}} \\ u_{t+1}^{\text{bal}} \\ \xi_t^{\text{bal}} \end{bmatrix}. \end{aligned} \quad (34)$$

To start the second step, I rewrite volume and the change in the spread as

$$vol_t = \alpha\beta E_t^* d_t - \alpha\omega E_t^* \Delta spr_{t+1} + u_{t+1}^{\text{vol}} \quad \text{and} \quad (35)$$

$$\Delta spr_{t+1} = -\beta\kappa^{\text{d}}(d_t + \xi_t^{\text{d}} - E_t d_t) + \beta vol_t - \beta(1 - \varphi^{\text{d}})(vol_t - E_t^* vol_t). \quad (36)$$

This implies that  $E_t \Delta spr_{t+1} = \beta E_t vol_t$  and hence  $E_t vol_t = \alpha\beta E_t d_t - \alpha\omega E_t \Delta spr_{t+1}$ , so combining these expressions gives

$$E_t vol_t = \frac{\alpha\beta}{1 + \alpha\beta\omega} E_t d_t \quad \text{and} \quad E_t \Delta spr_{t+1} = \beta \frac{\alpha\beta}{1 + \alpha\beta\omega} E_t d_t. \quad (37)$$

Equations (35) and (36) also imply that

$$E_t^* \Delta spr_{t+1} = -\beta\kappa^{\text{d}}(d_t + \xi_t^{\text{d}} - E_t d_t) + \beta E_t^* vol_t \quad \text{and}$$

$$E_t^* vol_t = \alpha\beta E_t^* d_t - \alpha\omega E_t^* \Delta spr_{t+1}.$$

Solving these equations gives

$$\begin{aligned} E_t^* vol_t &= \frac{\alpha\beta}{1 + \alpha\beta\omega} E_t d_t + \frac{\alpha\beta(1 + \omega)}{1 + \alpha\beta\omega} \kappa^{\text{d}}(d_t + \xi_t^{\text{d}} - E_t d_t) \quad \text{and} \\ E_t^* \Delta spr_{t+1} &= \beta \frac{\alpha\beta}{1 + \alpha\beta\omega} E_t d_t - \beta \left( \frac{1 - \alpha\beta}{1 + \alpha\beta\omega} \right) \kappa^{\text{d}}(d_t + \xi_t^{\text{d}} - E_t d_t). \end{aligned}$$

We can now find expressions for volume and the spread:

$$vol_t = (1 - \rho)E_t d_t + (1 + \omega)(1 - \rho)\kappa^d(d_t + \xi_t^d - E_t d_t) + u_{t+1}^{vol} \quad \text{and} \quad (38)$$

$$\Delta spr_{t+1} = \beta(1 - \rho)E_t d_t - \beta(1 - (1 + \omega)(1 - \rho))\kappa^d(d_t + \xi_t^d - E_t d_t) + \beta\varphi^d u_{t+1}^{vol}. \quad (39)$$

To complete the second step we solve for the estimated total depth and its associated estimation error. Equations (8), (12) and (14) imply that

$$E_{t+1}d_{t+1} = \rho E_t d_t + (\rho - \omega(1 - \rho))\kappa^d(d_t + \xi_t^d - E_t d_t) - \varphi^d u_{t+1}^{vol} \quad \text{and} \quad (40)$$

$$d_{t+1} - E_{t+1}d_{t+1} = (1 - \kappa^d)(d_t - E_t d_t) + u_{t+1}^d - \kappa^d \xi_t^d - (1 - \varphi^d)u_{t+1}^{vol}. \quad (41)$$

Simplifying these expressions gives

$$\begin{aligned} \begin{bmatrix} E_{t+1}d_{t+1} \\ d_{t+1} - E_{t+1}d_{t+1} \end{bmatrix} &= \begin{bmatrix} \rho & (1 - (1 + \omega)(1 - \rho))\kappa^d \\ 0 & 1 - \kappa^d \end{bmatrix} \begin{bmatrix} E_t d_t \\ d_t - E_t d_t \end{bmatrix} \\ &+ \begin{bmatrix} -\varphi^d & 0 & (1 - (1 + \omega)(1 - \rho))\kappa^d \\ -(1 - \varphi^d) & 1 & -\kappa^d \end{bmatrix} \begin{bmatrix} u_{t+1}^{vol} \\ u_{t+1}^d \\ \xi_t^d \end{bmatrix}. \end{aligned} \quad (42)$$

## A.2 Derivation of Equation (23)

First, we rewrite (22) as a difference equation in  $p_T$  and solve forward  $H$  periods. Applying the Law of Iterated Expectations to the resulting expression produces

$$P_T = E_T \sum_{i=0}^{H-1} (i_{T+i}^* - i_{T+i}) - \mathbb{E}_t^D \sum_{i=0}^{H-1} er_{T+i} + \bar{p}_T, \quad (43)$$

where  $\bar{p}_T \equiv E_T p_{T+H}$ . Next, we consider the implications of (43) for the daily depreciation rate,  $\Delta p_{T+1} = p_{T+1} - p_T$ . By definition, this rate equals the sum of expected depreciation rate,  $E_T \Delta p_{T+1}$ , and the forecast error  $p_{T+1} - E_T p_{T+1}$ ; components that can be directly computed from (43) as

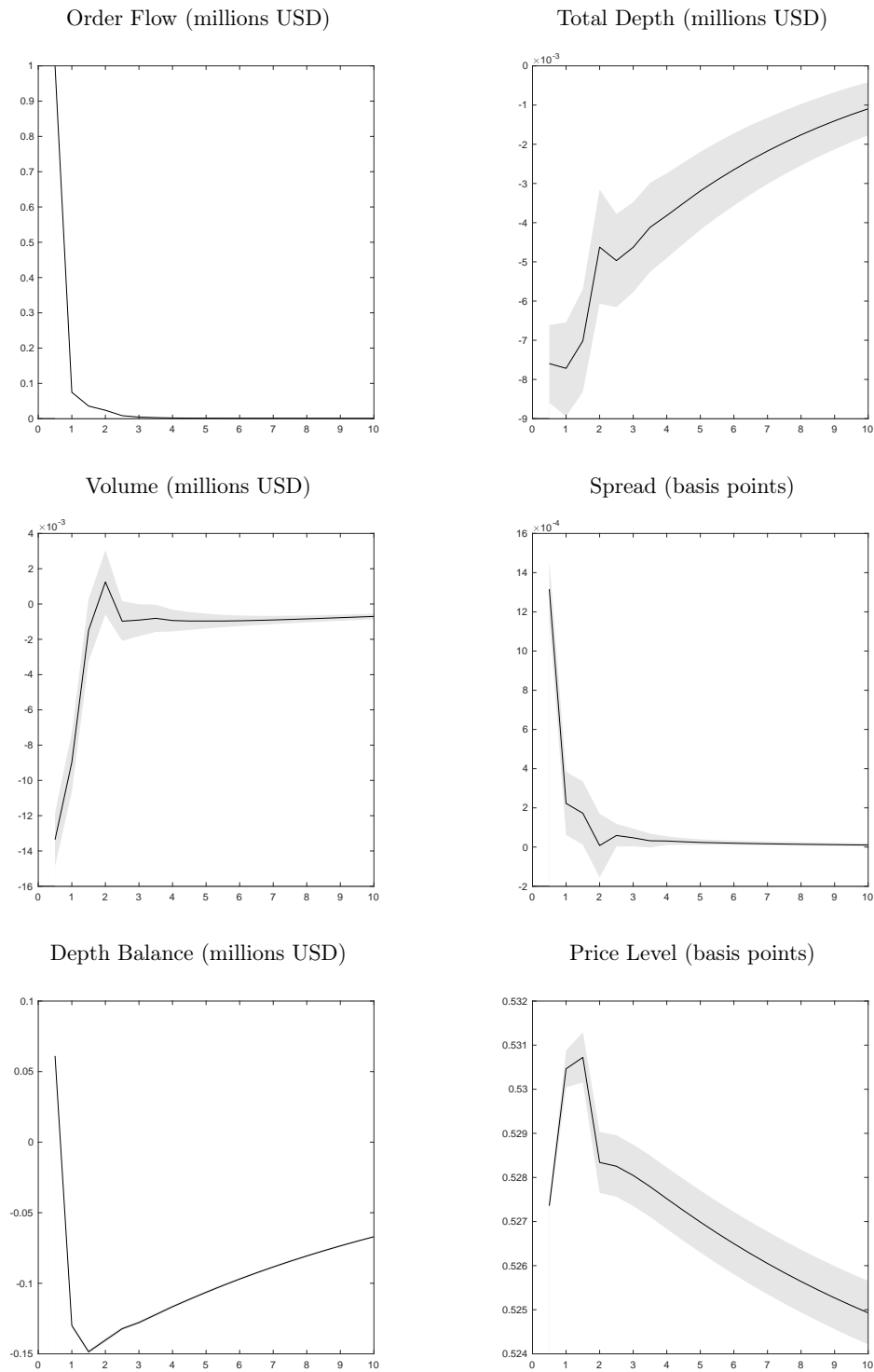
$$E_T \Delta p_{T+1} = i_T - i_T^* + er_T \quad \text{and} \quad (44a)$$

$$p_{T+1} - E_T p_{T+1} = -(E_{T+1} - E_T) \sum_{i=1}^H (i_{T+i} - i_{T+i}^*) - (E_{T+1} - E_T) \sum_{i=1}^H er_{T+i} + (\bar{p}_{T+1} - E_T \bar{p}_{T+1}). \quad (44b)$$

Substituting these expressions into the identity  $\Delta p_{T+1} = E_T \Delta p_{T+1} + p_{T+1} - E_T p_{T+1}$  produces equation (23).

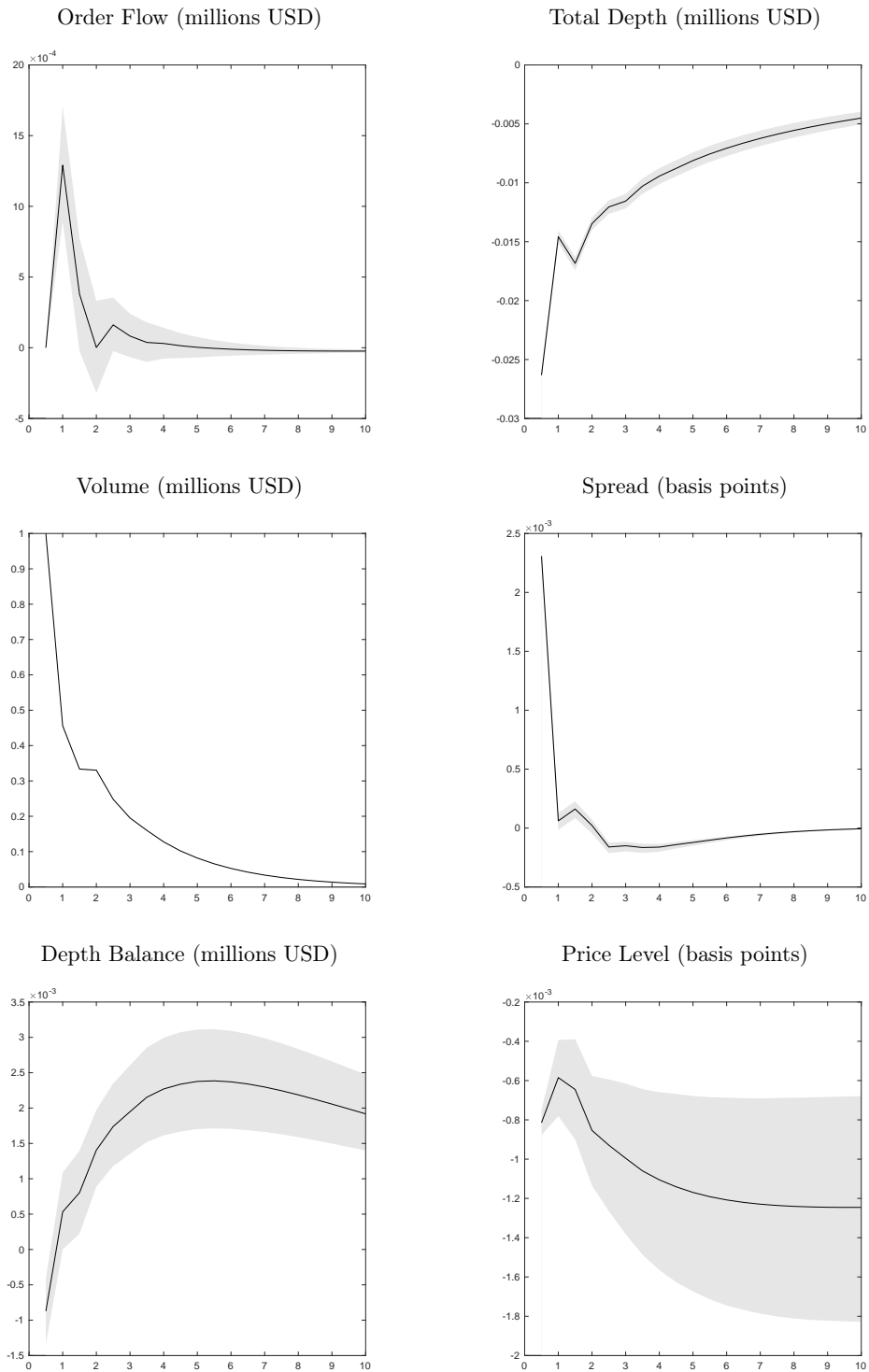
### A.3 VAR Impulse Responses

Figure 4: Impulse Responses to  $v_t^1$  Innovations



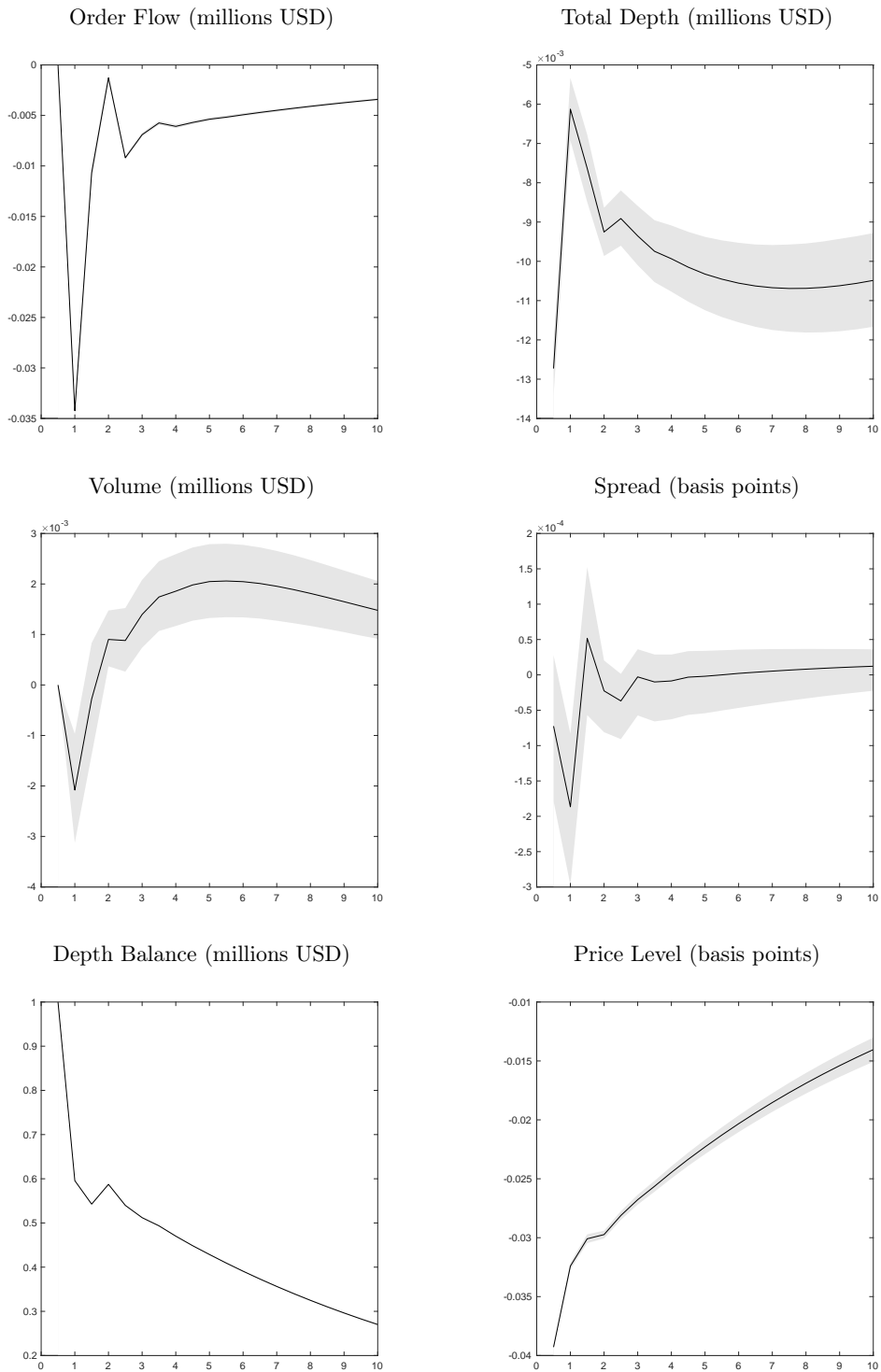
Notes: The plots show the impulse responses to a  $v_t^1$  innovation. The shaded bands around each response are 95% confidence bands computed from a bootstrap. The horizontal axis shows the number of minutes since the innovations.

**Figure 5: Impulse Responses to  $v_t^2$  Innovations**



Notes: The plots show the impulse responses a  $v_t^2$  innovation. The shaded bands around each response are 95% confidence bands computed from a bootstrap. The horizontal axis shows the number of minutes since the innovations.

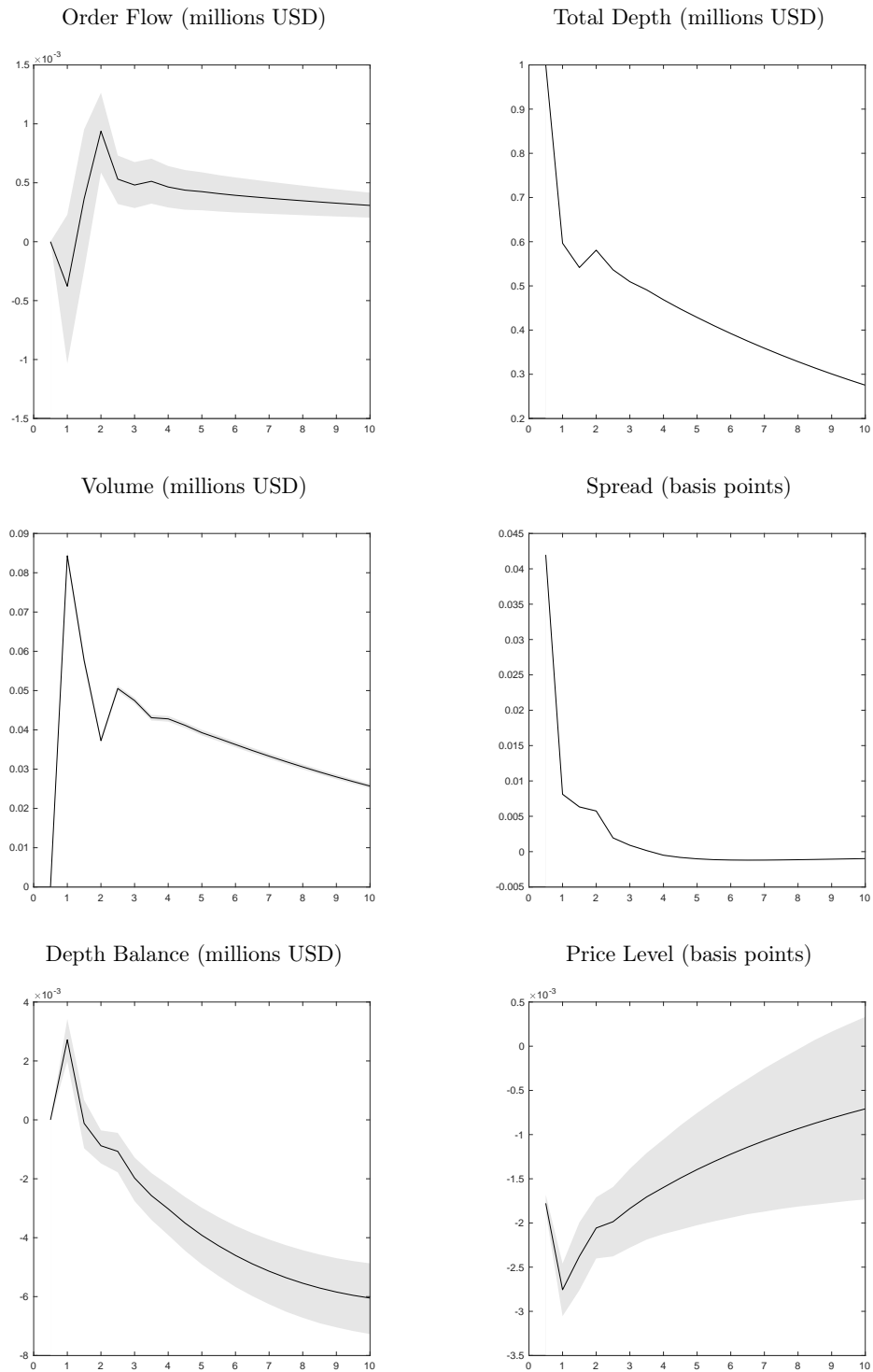
**Figure 6: Impulse Responses to  $v_t^3$  Innovations**



Notes: The plots show the impulse responses a  $v_t^3$  innovation. The shaded bands around each response are 95% confidence bands computed from a bootstrap. The horizontal axis shows the number of minutes since the innovations.

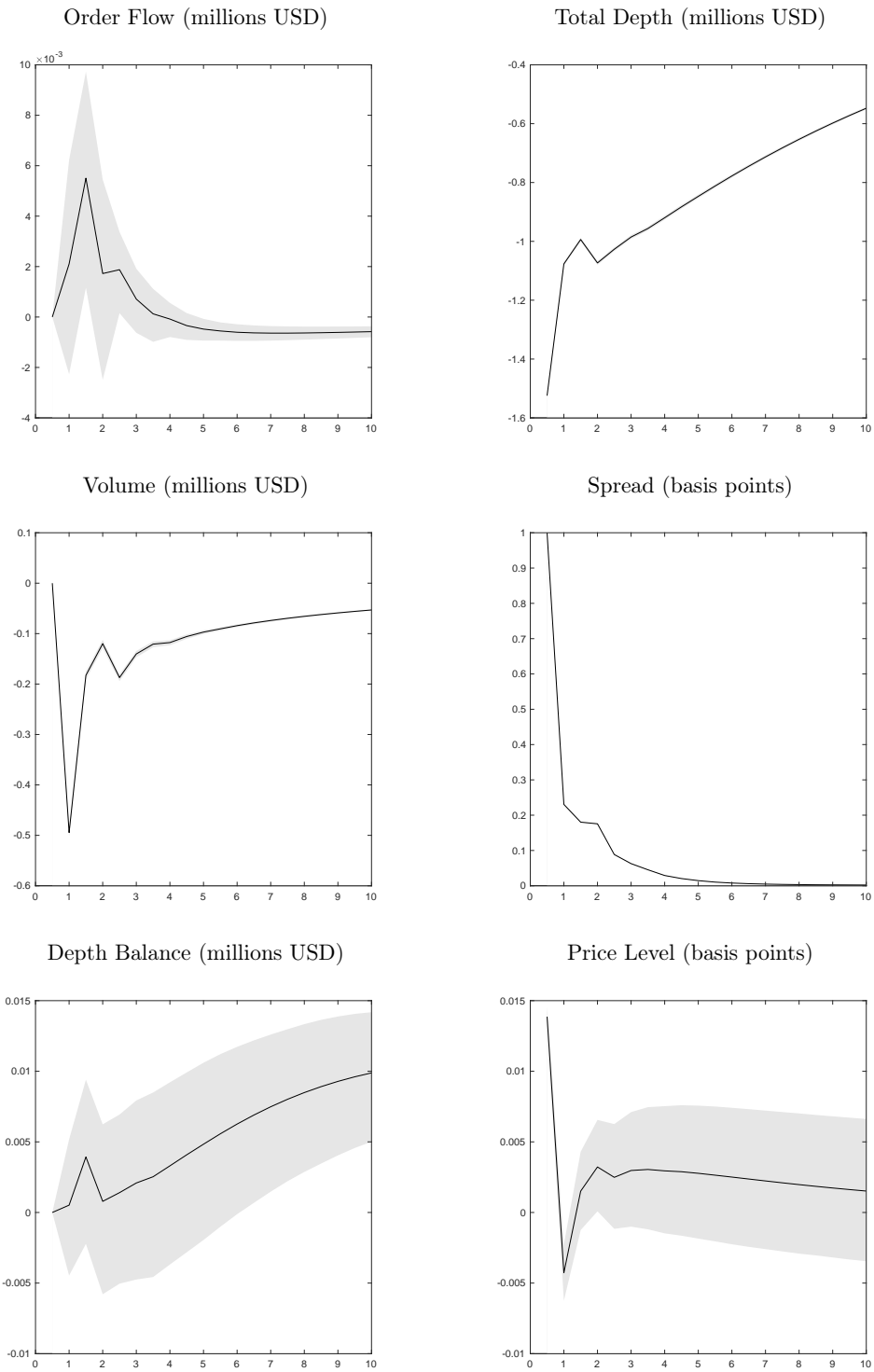


**Figure 7: Impulse Responses to  $v_t^4$  Innovations**



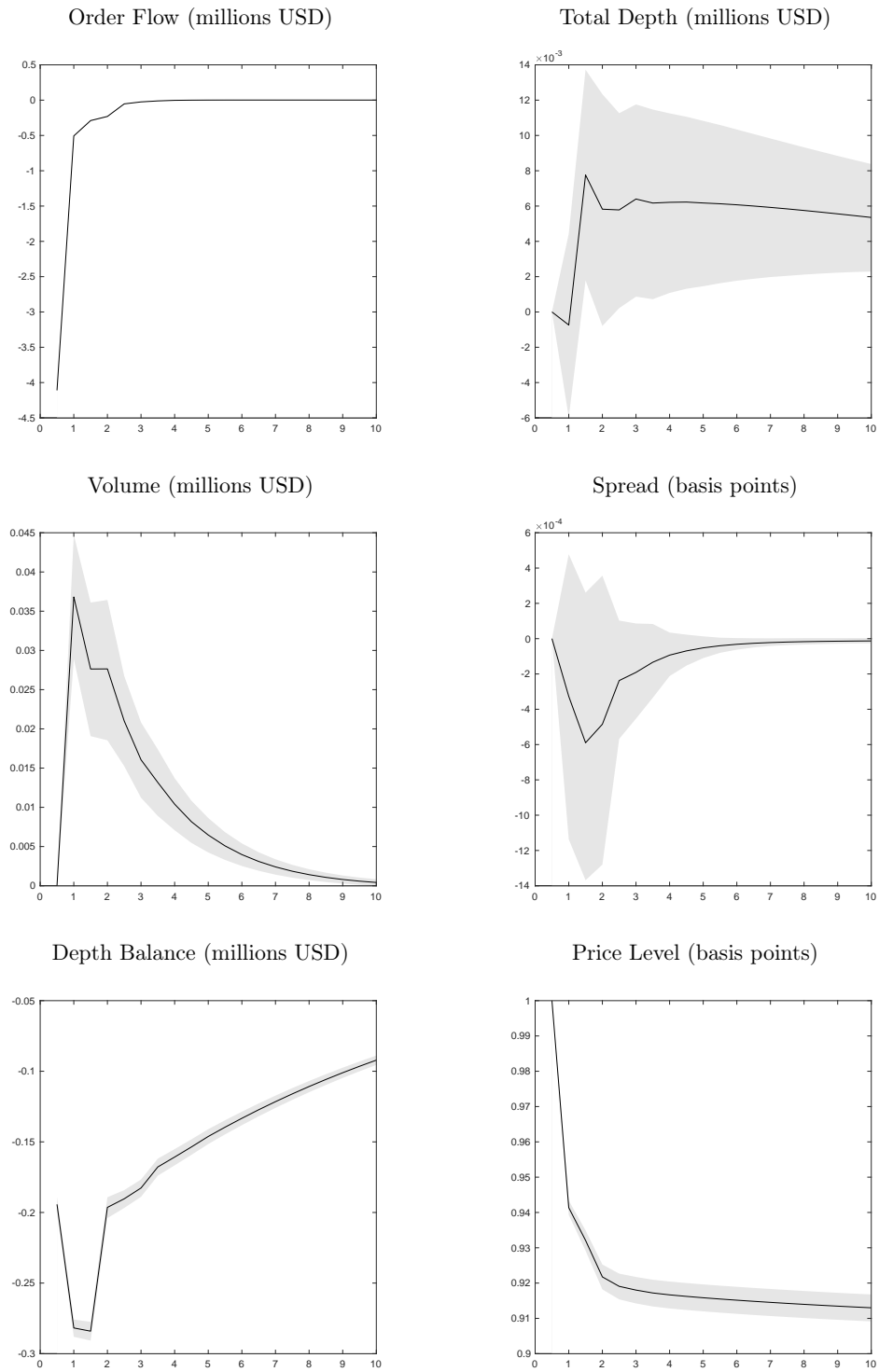
Notes: The plots show the impulse responses a  $v_t^4$  innovation. The shaded bands around each response are 95% confidence bands computed from a bootstrap. The horizontal axis shows the number of minutes since the innovations.

**Figure 8: Impulse Responses to  $v_t^5$  Innovations**



Notes: The plots show the impulse responses a  $v_t^5$  innovation. The shaded bands around each response are 95% confidence bands computed from a bootstrap. The horizontal axis shows the number of minutes since the innovations.

**Figure 9: Impulse Responses to  $v_t^6$  Innovations**



Notes: The plots show the impulse responses a  $v_t^6$  innovation. The shaded bands around each response are 95% confidence bands computed from a bootstrap. The horizontal axis shows the number of minutes since the innovations.

## A.4 Additional Regressions

**Table 8: Forecasting Order Flow Regressions**

Sample	Forecasting Variables	Forecast Variables					
		I: One Day Ahead			II: One Week Ahead		
		$flw$	$flw^1$	$flw^6$	$flw$	$flw^1$	$flw^6$
2003-2015							
	$flw^1$	0.131** (0.053)	0.000 (0.020)	0.066 (0.043)	0.219 (0.116)	0.039 (0.038)	0.172 (0.094)
	$flw^6$	0.297** (0.040)	0.070** (0.011)	0.256** (0.031)	0.206** (0.072)	0.049** (0.018)	0.166** (0.058)
	$\bar{R}^2$	0.066	0.023	0.073	0.043	0.023	0.045
2003-2006							
	$flw^1$	0.206** (0.056)	-0.042 (0.030)	0.041 (0.048)	0.276** (0.137)	0.010 (0.050)	0.209* (0.108)
	$flw^6$	0.257** (0.050)	0.082** (0.020)	0.229** (0.037)	0.326** (0.075)	0.081** (0.027)	0.306** (0.056)
	$\bar{R}^2$	0.052	0.015	0.053	0.090	0.026	0.119
2007-2011							
	$flw^1$	0.087 (0.098)	0.041 (0.032)	0.041 (0.080)	0.141 (0.208)	0.067 (0.064)	0.103 (0.169)
	$flw^6$	0.317** (0.052)	0.061** (0.015)	0.270** (0.041)	0.210** (0.101)	0.040 (0.026)	0.164** (0.081)
	$\bar{R}^2$	0.070	0.032	0.077	0.037	0.027	0.035
2012-2015							
	$flw^1$	0.115 (0.067)	-0.039 (0.040)	0.107** (0.045)	0.383** (0.133)	0.003 (0.060)	0.351** (0.105)
	$flw^6$	0.224** (0.044)	0.056** (0.026)	0.181** (0.032)	0.071 (0.080)	0.013 (0.033)	0.064 (0.061)
	$\bar{R}^2$	0.033	0.004	0.043	0.033	-0.001	0.052

Notes: The table reports estimates of forecasting regressions for order flow and the order flow components at the one day (panel I) and one week horizon (panel II). Forecasting variables are the daily order flow components in panel I and weekly order flow components in panel II. Regressions are estimated by OLS at the daily frequency. Heteroskedastic-consistent standard errors are reported in parenthesis below each of the parameter estimates. Their standard errors also account for the MA(4) process induced by the overlapping forecast errors in the panel II regressions. “\*” and “\*\*” denote statistical significance at the 5 and 1 percent levels, respectively.