Online Appendix for 'The Widening of Cross-Currency Basis: When Increased FX Swap Demand Meets Limits of Arbitrage'

Nadav Ben Zeev*
Ben-Gurion University of the Negev

Daniel Nathan†
University of Pennsylvania and Bank of Israel

July 22, 2024

Abstract

This online appendix consists of the following two appendices: an appendix presenting the details of the model that underlies the paper's theoretical motivation; an appendix detailing the estimation procedure and results for the local projection extension of the baseline econometric model; and an appendix presenting the raw coefficient estimate results corresponding to Tables 7, 8, 9, 10, and 12 from the main text.

*Department of Economics, Ben-Gurion University of the Negev, Beer-Sheva, Israel. *E-mail:* nadavbz@bgu.ac.il.

†Finance Department, The Wharton School, University of Pennsylvania, Philadelphia, United States, and the Research Department, Bank of Israel, Jerusalem, Israel. *E-mail:* nathad@wharton.upenn.edu.

Appendix A Theoretical Motivation

In what follows we lay out a simple structural framework which is meant to fix ideas and form a suitable conceptual base for our paper's empirical analysis. Understanding the drivers of CIP deviations is tantamount to understanding the workings of the FX swap market (see Du and Schreger (2022) and references therein). Accordingly, the framework we use is a partial equilibrium of the FX swap market that builds heavily on Liao and Zhang (2020). The model consists of two time periods (t and $t+1$) and two agents. The first agent is a risk-neutral arbitrageur who supplies FX swaps. The second is a risk-averse local institutional investor (II) who demands FX swaps to obtain FX-risk-free foreign currency funding. The use of this foreign currency funding is for the local II to increase its (hedged) exposure to foreign assets.

We start our depiction of the model with a presentation of the supply side of the FX swap market by presenting the arbitrageur's supply of FX swaps. We then show demand for FX swaps by the local II. We end the section by defining equilibrium and presenting the model's main prediction.

A.1 Supply of FX Swaps

General Setting. There is a risk-neutral arbitrageur that represents the supply side of the FX swap market. The arbitrageur's aim is to profit from the local II by creating a synthetic forward rate that is cheaper than the market's observed forward rate. The arbitrageur's trade can be broken down into two parts. First, it buys spot $Q_{t,ARB}S_t$ local currency units and sells spot $Q_{t,ARB}$ foreign currency units in period t, conducting this trade entirely with the local II. Second, it sells forward $Q_{t,ARB}S_t(1+i_{t+1,L})$ local currency units at forward rate $F_{t,t+1}$, where $Q_{t,ARB}S_t$ is the local currency amount sold forward to the local II in the second leg of the associated arbitrageur-local II FX swap trade and $Q_{t,ARB}S_t i_{t+1,L}$ represents the interest related amount sold forward in an outright forward trade the arbitrageur conducts with some (unmodeled) broker-dealer institution.

The rationale for the second part of the trade can be explained as follows. Using its predetermined arbitrage capital, the arbitrageur conducts CIP arbitrage as well as other arbitrage trades (whose depiction is deferred for now). Given its role as arbitrageur, and since FX swaps

trades do not perfectly align with CIP arbitrage as they exclude the interest proceeds element,[1] the arbitrageur additionally sells forward $Q_{t,ARB} S_t i_{t+1,L}$ local currency units, where $Q_{t,ARB}$ is the arbitrageur's FX swap supply (in foreign currency units) and $i_{t+1,L}$ is the local risk-free interest rates, respectively. (While left unmodeled, the counterparty to this interest proceeds forward trade can be thought of as a broker-dealer institution.)

The arbitrageur's level of predetermined arbitrage capital will be used below as the model's LOA measure in the sense that a lower such capital level implies greater LOA. While this LOA representation constitutes a reduced-form encapsulation of LOA, we view it as the most natural way to represent LOA in our structural setting.

We assume that the arbitrageur can borrow foreign currency frictionlessly in the cash market at interest rate $i_{t+1,W}$ and hence has no constraints on its funding of foreign currency. (I.e., $i_{t+1,W}$ represents the opportunity cost of arbitrageur's FX swap trade. In our setting it is viewed as the effective cost of the FX swap trade as we assume the arbitrageur funds this trade by borrowing the required funds in the cash market.) However, we assume that it faces frictions in the FX swap market, as we now turn to explain.

Haircut. Following Ivashina et al. (2015), we assume that a haircut (initial margin) is applied to the arbitrageur's FX swap trade in the amount of $\kappa Q_{t,ARB}$. That is, the arbitrageur's FX swap trade requires it to incur a linear haircut-induced cost through the depositing of share κ of its swap position to the local II.[2]

Arbitrageur's Alternative Arbitrage Activity. By allocating $\kappa Q_{t,ARB}$ for CIP arbitrage, the arbitrageur has to take these funds away from its pre-determined arbitrage capital A_t. In other words, $A_t - \kappa Q_{t,ARB}$ represents the arbitrageur's available capital for another (non-CIP) arbitrage activity (e.g., fixed income arbitrage). Following Ivashina et al. (2015), this other arbitrage activity has a net *concave* return given by $G(A_t - \kappa Q_{t,ARB})$, where $G(:) > 0$, $G'(:) > 0$, $G''(:) < 0$, and

[1] While this element is necessary for conducting CIP arbitrage, it is noteworthy that it is also very small (relative to the principal) given that FX swap trades' maturities are usually short (Schrimpf and Sushko (2019)).

[2] For simplicity, we abstract from the opposing haircut facing the local II.

$G'''(:) > 0$. [3]

These assumptions on $G(:)$ are met by standard production/revenue functions, including the logarithmic specification used in Ivashina et al. (2015). More generally, considering the commonly used positively homogenous production/revenue functions, it is straightforward to show that concavity ($G''(:) < 0$) in fact implies a positive third derivative ($G'''(:) > 0$) as the latter condition requires a returns to scale that is lower than 2 while the former implies a returns to scale that is lower than 1 (i.e., decreasing returns to scale). This is an important observation because evidence from the literature on bank investment returns (see Zhu (2008) and references therein) and the literature on mutual fund investment returns (McLemore (2019)) supports the notion of decreasing return to scale for financial institutions' investments.

Arbitrageur's Profit Maximization. We are now in position to write the arbitrageur's profit from its arbitrage activity as

$$Q_{t,ARB}\frac{S_t}{F_{t,t+1}}(1+i_{t+1,L}) - Q_{t,ARB}(1+i_{t+1,W}) + G(A_t - \kappa Q_{t,ARB}). \tag{A.1}$$

The FOC that results from maximizing the profit from Equation (A.1) with respect to $Q_{t,ARB}$ is

$$b_t \equiv 1 + i_{t+1,W} - \frac{S_t}{F_{t,t+1}}(1+i_{t+1,L}) = -\kappa G'(A_t - \kappa Q_{t,ARB}), \tag{A.2}$$

where b_t is the cross-currency basis (defined in accordance with the literature) and $\frac{S_t}{F_{t,t+1}}(1+i_{t+1,L})$ represents the synthetic, CIP-implied foreign (gross) risk-free interest rate which is clearly higher than the cash market one owing to the haircut-induced cost. In other words, Equation (A.2) implies a negative cross-currency basis b_t that is caused by the swap trade's haircut-induced friction.

Relation between $Q_{t,ARB}$ and $-b_t$. Minus of the cross-currency basis (i.e., $-b_t$) from FOC (A.2) is the arbitrageur's marginal profit from increasing its FX swap position. As such, the minus

[3] The concavity of G is consistent with the limits-to-arbitrage notion from Shleifer and Vishny (1997). For internal consistency between such arbitrage limits existing across all of the arbitrageur's arbitrage activities, we could also have assumed a *convex* haircut-induced cost as in Liao and Zhang (2020) which seems more consistent with such limits than our linear haircut assumption. Such modeling choice does not change our model's main prediction and hence, for simplicity, we stick to the linear haircut modeling approach from Ivashina et al. (2015).

of the cross-currency basis can also be economically viewed as the price of the FX swap. Accordingly, it is therefore reasonable to expect that the arbitrageur's supply of FX swaps increases in $-b_t$. To show this formally, we differentiate $-b_t$ from FOC (A.2) with respect to $Q_{t,ARB}$:

$$\frac{\partial(-b_t)}{\partial Q_{t,ARB}} = -\kappa^2 G''(A_t - \kappa Q_{t,ARB}) > 0, \tag{A.3}$$

where the positive sign of Equation (A.3) comes from the assumed concavity of net return function $G(A_t - \kappa Q_{t,ARB})$. Given the interpretation of $-b_t$ as FX swap price, Equation (A.3) delivers the standard result of an upward-sloping supply curve: higher price (marginal profit) of FX swaps induces the arbitrageur to supply more such swaps. Moreover, we can show that the slope of the arbitrageur's FX swap supply curve flattens (steepens) when initial arbitrage capital is higher (lower) by differentiating Equation (A.3) with respect to A_t:

$$\frac{\partial^2(-b_t)}{\partial Q_{t,ARB}\partial A_t} = -\kappa^2 G'''(A_t - \kappa Q_{t,ARB}) < 0. \tag{A.4}$$

This result clearly follows from the arguably weak assumption of $G(:)$'s positive third derivative (see related discussion on this assumption on Page 3). I.e., more (less) initial arbitrage capital induces less (more) rigidity in the willingness of the arbitrageur to supply FX swaps. Result (A.4) lies at the heart of our paper.

A.2 Demand for FX Swaps

General Setting. We assume a risk-averse local II that borrows in the swap market $Q_{t,II}$ foreign currency units for the purchase of foreign assets whose expected rate of return is denoted by $\mathbb{E}_t i_{t+1,FA}$, where $\mathbb{E}_t$ is the expectation operator conditional on period t information. (The $i_{t+1,FA}$ return variable can be thought of as some weighted average of returns of foreign stocks, bonds, and loans.) Specifically, the local II enters an FX swap with the arbitrageur of size $Q_{t,II}$. In the first leg of the trade the local II sells $Q_{t,II}S_t$ local currency spot units and buys $Q_{t,II}$ foreign currency units. And in the second leg, which takes place in period $t+1$, the local II buys $Q_{t,II}S_t$ local currency units at forward rate $F_{t,t+1}$ and sells $\frac{Q_{t,II}S_t}{F_{t,t+1}}$ foreign currency units. We abstract from the haircut that the local II realistically faces in this swap trade as well as from its non-swap-related investments. Adding these elements would complicate the exposition without affecting the main prediction of our model.

Expectation and Variance of Local II's Profit. We can write the local II's next period's expected profit (in foreign currency terms) from its swap-related foreign investment, which we assume to be positive and denote by $\mathbb{E}_t\Pi_{t+1,II}$, as

$$\mathbb{E}_t\Pi_{t+1,II} = Q_{t,II}(1 + \mathbb{E}_t i_{t+1,FA}) - Q_{t,II}\frac{S_t}{F_{t,t+1}}. \tag{A.5}$$

We can use the definition of cross-currency basis from Equation (A.2) to write Equation (A.5) equivalently as

$$\mathbb{E}_t\Pi_{t+1,II} = Q_{t,II}(1 + \mathbb{E}_t i_{t+1,FA}) - Q_{t,II}\left(\frac{1 + i_{t+1,W} - b_t}{1 + i_{t+1,L}}\right). \tag{A.6}$$

And the variance of local II's profit ($\mathbb{V}_t\Pi_{t+1,II}$) can be written as $\mathbb{V}_t\Pi_{t+1,II} = Q^2_{t,II}\mathbb{V}_t(i_{t+1,FA})$, where $\mathbb{V}_t$ is the variance operator conditional on period t information.

Mean-Variance Optimization Problem. We assume the local II chooses its demand for FX swaps $Q_{t,II}$ so as to maximize

$$\mathbb{E}_t\Pi_{t+1,II} - \frac{e^{\epsilon_t}}{2}\mathbb{V}_t\Pi_{t+1,II} = Q_{t,II}(1 + \mathbb{E}_t i_{t+1,FA}) - Q_{t,II}\left(\frac{1 + i_{t+1,W} - b_t}{1 + i_{t+1,L}}\right) - \frac{e^{\epsilon_t}}{2}Q^2_{t,II}\mathbb{V}_t(i_{t+1,FA}), \tag{A.7}$$

where ϵ_t represents an FX swap demand white noise shock which in turn determines the level of local II's risk aversion with respect to swap-related foreign investment. Importantly, as formally shown below, a positive (negative) ϵ_t induces a leftward (rightward) shift in the demand for FX swaps. More generally, when one considers the alternative local investment opportunities facing the local II, such shocks essentially represent exogenous shifts in the local II's geographical investment preferences.

The FOC that results from maximizing the objective function from Equation (A.7) with respect to $Q_{t,II}$ is

$$Q_{t,II} = \frac{1 + \mathbb{E}_t i_{t+1,FA}}{e^{\epsilon_t}\mathbb{V}_t(i_{t+1,FA})} - \frac{1 + i_{t+1,W} - b_t}{(1 + i_{t+1,L})e^{\epsilon_t}\mathbb{V}_t(i_{t+1,FA})}. \tag{A.8}$$

Equation (A.8) essentially represents local II's demand for FX swaps.

Relation between $Q_{t,II}$ and $-b_t$. In the previous section we interpreted $-b_t$ as the price of FX swaps. As such, we should expect to have a negative relation between this price and demand for FX swaps. To show this negative relation (i.e., a downward sloping FX swap demand curve), let us differentiate Equation (A.8) with respect to $-b_t$:

$$\frac{\partial Q_{t,II}}{\partial(-b_t)} = -\frac{1}{(1+i_{t+1,L})e^{\epsilon_t}\mathbb{V}_t(i_{t+1,FA})} < 0. \tag{A.9}$$

Relation between $Q_{t,II}$ and ϵ_t. We argued above that a positive (negative) realization for ϵ_t represents a leftward (rightward) shift in local II's FX swap demand. To show this formally, let us differentiate Equation (A.8) with respect to e^{ϵ_t}:

$$\frac{\partial Q_{t,II}}{\partial e^{\epsilon_t}} = -\frac{1+\mathbb{E}_t i_{t+1,FA}}{e^{2\epsilon_t}[\mathbb{V}_t(i_{t+1,FA})]^2} + \frac{1+i_{t+1,W}-b_t}{e^{2\epsilon_t}[\mathbb{V}_t(i_{t+1,FA})]^2(1+i_{t+1,L})} < 0, \tag{A.10}$$

where the negative sign of Equation (A.10) comes from the fact that local II's expected profit, $1+\mathbb{E}_t i_{t+1,FA} - \left(\frac{1+i_{t+1,W}-b_t}{1+i_{t+1,L}}\right)$, is assumed to be positive.

A.3 Model Equilibrium

We define equilibrium in the FX swap market as the equality $Q_{t,II} = Q_{t,ARB} = Q_t$, where Q_t denotes the equilibrium level of FX swap flows. The latter equilibrium equation, when substituted into FOCs (A.2) and (A.8) produce two equations in two unknowns b_t and Q_t. (A proof that relies on a fixed-point argument for the existence and uniqueness of a solution to this demand-supply equation system is available upon request from the authors.) We can use our previous results on the nature of the FX swap supply and demand curves to deduce the main prediction of our model.

The A_t-Dependent Relation Between ϵ_t and b_t. Consider our model's FX demand-supply framework in the space of $-b_t$ and Q_t. Equation (A.2) defines an upward-sloping FX swap supply curve whose slope becomes steeper with a lower A_t. Equation (A.8) defines a downward-sloping FX swap demand curve which shifts rightward in response to a negative realization of swap demand shock ϵ_t. In equilibrium, such favorable swap demand shock is predicted to produce an

increase in $-b_t$ (i.e., a widening of the basis) which depends on the level of the arbitrageur's initial arbitrage capital A_t: the lower (higher) this capital is, the stronger (weaker) the widening effect of the demand shock.

To see this relation formally, we take three steps. First, we substitute Equation (A.8) into Equation (A.2) (after substituting into both equations the equilibrium condition $Q_{t,II} = Q_{t,ARB} = Q_t$) to obtain the following equilibrium equation for b_t:

$$b_t = -\kappa G'\left(A_t - \kappa\left(\frac{1+\mathbb{E}_t i_{t+1,FA}}{e^{\epsilon_t}\mathbb{V}_t(i_{t+1,FA})} - \frac{1+i_{t+1,W}-b_t}{(1+i_{t+1,L})e^{\epsilon_t}\mathbb{V}_t(i_{t+1,FA})}\right)\right). \tag{A.11}$$

Second, we implicitly differentiate Equation (A.11) with respect to e^{ϵ_t} to obtain the effect of the latter on b_t:[4]

$$\frac{\partial b_t}{\partial e^{\epsilon_t}} = \frac{\kappa^2 G''(:)\frac{\partial Q_{t,II}}{\partial e^{\epsilon_t}}}{1-\kappa G''(:)\frac{\partial Q_{t,II}}{\partial b_t}} > 0. \tag{A.12}$$

The positive sign of Equation (A.12) relies on the assumed concavity of G and the derived negative signs of $\frac{\partial Q_{t,II}}{\partial(-b_t)}$ and $\frac{\partial Q_{t,II}}{\partial e^{\epsilon_t}}$ from Equations (A.9) and (A.10), respectively. Third, we differentiate Equation (A.12) with respect to A_t:

$$\frac{\partial^2 b_t}{\partial e^{\epsilon_t}\partial A_t} = \frac{\kappa^2 G'''(:)\frac{\partial Q_{t,II}}{\partial e^{\epsilon_t}}(1-\kappa G''(:)\frac{\partial Q_{t,II}}{\partial b_t}) + \kappa^2 G''(:)\frac{\partial Q_{t,II}}{\partial e^{\epsilon_t}}G'''(:)\frac{\partial Q_{t,II}}{\partial b_t}}{\left(1-\kappa G''(:)\frac{\partial Q_{t,II}}{\partial b_t}\right)^2} = \tag{A.13}$$

$$\frac{\kappa^2 G'''(:)\frac{\partial^2 b_t}{\partial e^{\epsilon_t}}\left(1-G''(:)\frac{\partial Q_{t,II}}{\partial b_t}(\kappa^2-\kappa)\right)}{\left(1-\kappa G''(:)\frac{\partial Q_{t,II}}{\partial b_t}\right)^2} < 0.$$

The negative sign of Equation (A.13) relies on the assumed concavity of G, its assumed positive third derivative, the fact that $\kappa < 1$, and the derived negative signs of $\frac{\partial Q_{t,II}}{\partial(-b_t)}$ and $\frac{\partial Q_{t,II}}{\partial e^{\epsilon_t}}$ from Equations (A.9) and (A.10), respectively. Equations (A.12) and (A.13) formally demonstrate that a negative realization for ϵ_t (i.e., a rightward shift in FX swap demand) is predicted to generate a stronger widening of the basis (i.e., a larger decline in b_t) if the initial value of A_t is lower (i.e., if LOA are greater).

[4]To streamline the remaining two derivations' exposition, which is otherwise quite cumbersome, we avoid writing out the argument in G as well as the explicit expressions from $\frac{\partial Q_{t,II}}{\partial(-b_t)}$ and $\frac{\partial Q_{t,II}}{\partial e^{\epsilon_t}}$ from Equations (A.9) and (A.10). The signs of these expressions are sufficient for our purposes in these two derivations.

This prediction has strong economic intuition given that lower A_t, by limiting the availability of funds for arbitrageurs' arbitrage activity and thus inducing greater LOA, should make their FX swap supply more rigid and hence make the basis (b_t) more responsive to a rightward shift in FX swap demand. The A_t-dependent FX swap demand channel embodied by Equation (A.13) can also be equivalently referred to as the LOA-dependent FX swap demand channel (as done in the previous sections as well as hereafter), which is the central object of study of the paper.

Appendix B Local Projection Extension

While our paper focused on the *impact* LOA-dependent effect of demand shocks - a natural choice given this object's obvious centrality in our analysis -, it seems worthwhile to also explore the dynamic nature of the LOA-dependent channel of FX swap demand as this would inform us about the magnitude of this channel's persistence. Toward this end, and having established our GIV-with-controls approach as our preferred method for estimation of the LOA-dependent channel of FX swap demand, we present here a dynamic extension of the latter approach which allows for estimation of impulse responses and forecast error variance (FEV) contributions for our GIV shock.

We implement this extension by jointly estimating a local projection regression counterpart to Equation (8) from the text with Equations (4), (5), and (6) in a way that accounts for estimation uncertainty surrounding all of these equations. As such, in addition to informing us about the dynamic nature of the LOA-dependent channel of FX swap demand studied in this paper, this extension also serves to alleviate the concern that our impact-based results are sensitive to abstracting from estimation uncertainty surrounding the extraction of the GIV shock. We use a Bayesian estimation and inference procedure as it provides a convenient numerical way to produce confidence intervals that account for estimation uncertainty in each of the equations underlying our GIV estimation procedure.[5]

[5]The Bayesian approach we take is in the spirit of a long tradition in the literature on impulse response estimation (see, e.g., Del Negro and Schorfheide (2011)) that has recently also caught on in the local projections literature (see, e.g, Miranda-Agrippino and Ricco (2021) and Ben Zeev (2023)).

B.1 Econometric Model

The model we estimate is

$$\Delta SP_{i,t} = \mathbb{C}'\gamma_{i,L} + LOA_{t-1}\mathbb{C}'\gamma_{i,I} + v_{i,t}, \tag{B.1}$$

$$\hat{v}_{i,t} = \epsilon_{i,t} + \zeta_i LOA_{t-1}\epsilon_{i,t} + \eta_{i,t}, \tag{B.2}$$

$$\epsilon_{i,t} = \mu_{i,t}, \tag{B.3}$$

$$b_{t+h-1} - b_{t-1} = \alpha_h + \Xi_{L,h}\varrho_{GIV,t} + \Xi_{I,h}LOA_{t-1}\varrho_{GIV,t} + \beta_h LOA_{t-1} + u_{t+h-1}, \tag{B.4}$$

where Equations (B.1)-(B.3) are identical to Equations (4)-(6) from the text and Equation (B.4) is the local projection regression counterpart to Equation (4) from the text with $h = 1, 2, ..., 10$ denoting the rolling horizon and Equation (B.4) for $h = 1$ (i.e., the impact horizon) being identical to Equation (4) from the text. (We consider 10 trading days). Accounting for non-trading days in the USD/NIS FX swap market, these 10 trading days correspond to roughly 20 calendar days. Following Gabaix and Koijen (2023), we define the GIV shock (denoted by $\varrho_{GIV,t}$) as the difference between the size-weighted- and equal-weighted-average of the estimated idiosyncratic shocks, i.e., $\varrho_{GIV,t} = \sum_{i=1}^{13} \hat{\epsilon}_{i,t} w_i - \sum_{i=1}^{13} \hat{\epsilon}_{i,t} \frac{1}{13}$, where w_i is II i's share of swap flows' average volume in the sum of IIs' average volumes. The impulse responses in the LOA and no LOA states are $\Xi_{L,h} + 2\Xi_{I,h}$ and $\Xi_{L,h} - 1.4\Xi_{I,h}$, respectively, providing us in turn estimates of the LOA-dependent effects of the GIV shock over roughly 20 calendar days.

B.2 Posterior Distribution of Parameters

We estimate Equations (B.1)-(B.4) jointly by applying the Bayesian estimation algorithm for strong block-recursive structure put forward by Zha (1999) for block-recursive VARs, where the likelihood function is broken into the different recursive blocks. In our case, we only have three blocks, where the first consists of Equation (B.1), the second consists of Equations B.2 and (B.3), and the third contains Equation (B.4). As shown in Zha (1999), this kind of block separation along with the standard assumption of a normal-inverse Wishart conjugate prior structure leads to a normal-inverse Wishart posterior distribution for the block-recursive equation parameters.

To account for temporal correlations of the error term in Equation (B.4), we apply a Newey-West correction to the standard errors within our Bayesian estimation procedure. In doing so

we accord with the reasoning from Miranda-Agrippino and Ricco (2021), who estimate a hybrid VAR-local-projections model and follow the suggestion from Müller (2013) to increase estimation precision in the presence of a misspecified likelihood function (as in our and their setting) by replacing the original posterior's covariance matrix with an appropriately modified one. Moreover, given the high-frequency nature of our data and the general tendency of impulse responses from local projections to exhibit jaggedness, we apply a suitably modified version of the smoothing procedure from Plagborg-Møller (2016) to our estimated raw impulse responses.

Given the block-recursiveness of the econometric model (Equations (B.1)-(B.4)), we present the estimation algorithm separately for the three blocks underlying these equations.

B.3 Equation (B.1)

Companion Form of Specification. Keeping with the notation from Section 5.2.1 of the paper, Equation (B.1) can be written in companion form as follows:

$$Y_i = X_i B_i + v_i, \tag{B.5}$$

where i indexes IIs and $Y_i = [\Delta SP_{i,1}, ..., \Delta SP_{i,T}]'$, $X_i = [X_{i,1}, ..., X_{i,T}]'$ where $X_{i,t}$ is a vector containing the RHS variables of Equation (B.1) and B_i is a vector that includes $X_{i,t}$'s corresponding coefficients; $v_i = [v_{i,1}, ..., v_{i,T}]'$ (with T being the time dimension of the sample and p_i being the number of lags); and $\sigma^2_{i,v}$ is the variance of the reduced form innovation series $v_{i,t}$ from that equation.

Specification of Uninformative Prior. We follow the conventional approach of specifying a normal-inverse Wishart prior distribution for the reduced-form parameters:[6]

$$vec(B_i) \mid \sigma^2_{i,v} \sim N(vec(\bar{B}_0), \sigma^2_{i,v} \otimes N_0^{-1}), \tag{B.6}$$

$$\sigma^2_{i,\epsilon} \sim IW_1(m_0 S_0, m_0), \tag{B.7}$$

where N_0 is a KxK positive definite matrix (K is the number of parameters for the RHS of Equation (B.1)), S_0 is a scalar, and $m_o > 0$. As shown by Uhlig (1994), the latter prior implies the following

[6]Since Equation (B.1) is univariate, the assumed inverse-Wishart distribution for $\sigma^2_{i,v}$ is denoted by IW_1, where the subscript 1 represents the univariate nature of Equation (B.1) and its corresponding univariate inverse-Wishart distribution.

posterior distribution:

$$vec(B_i) \mid \sigma^2_{i,v} \sim N(vec(\bar{B}_{i,T}), \sigma^2_{i,v} \otimes N^{-1}_{i,T}), \tag{B.8}$$

$$\sigma^2_{i,v} \sim IW_1(m_T S_{i,T}, m_T), \tag{B.9}$$

where $m_T = T + m_0$, $N_{i,T} = N_0 + X_i'X_i$, $\bar{B}_{i,T} = N^{-1}_{i,T}(N_0\bar{B}_0 + X_i'X_i\hat{B}_i)$,
$S_{i,T} = \frac{m_0}{m_T}S_0 + \frac{T}{m_T}\hat{\sigma^2_{i,v}} + \frac{1}{m_T}(\hat{B}_I - \bar{B}_0)'N_0N^{-1}_{i,T}X_i'X_i(\hat{B}_i - \bar{B}_0)$, $\hat{B}_i = (X_i'X_i)^{-1}X_i'Y_i$,
and $\hat{\sigma^2_{i,v}} = (Y_i - X_i\hat{B}_i)'(Y_i - X_i\hat{B}_i)/T$.

We follow the literature and use a weak prior, i.e., $m_0 = 0$, $N_0 = 0$, and arbitrary S_0 and $\bar{B}_0$. This implies that the prior distribution is proportional to $\sigma^2_{i,v}$ and that $m_T = T$, $S_{i,T} = \hat{\sigma^2_{i,v}}$, $\bar{B}_{i,T} = \hat{B}_i$, and $N_{i,T} = X_i'X_i$.

Posterior Simulator. In light of the above-described prior formulation, the posterior simulator for B_i and $\sigma^2_{i,v}$ can be described as follows:

1. Draw $\sigma^2_{i,v}$ from an $IW_1(T\hat{\sigma^2_{i,v}}, T)$ distribution.
2. Draw B_i from the conditional distribution $MN(\hat{B}_i, \sigma^2_{i,v} \otimes (X_i'X_i)^{-1})$.
3. Repeat steps 1 and 2 a large number of times and collect the drawn B_i's and $\sigma^2_{i,v}$'s.[7]

Once we have these draws at hand, we turn to estimate the second block of equations (Equations B.2 and (B.3)), which we now describe.

B.4 Equations (B.2) and (B.3)

Denoting the estimated residual of Equation (B.1) by $\hat{v}_{i,t}$, we formulate the relation between $\hat{v}_{i,t}$ and the sought-after true idiosyncratic demand shock $\epsilon_{i,t}$ as a time-varying state-space model:

$$\hat{v}_{i,t} = \epsilon_{i,t} + \zeta_i LOA_{t-1}\epsilon_{i,t} + \eta_{i,t}, \tag{B.10}$$

$$\epsilon_{i,t} = \mu_{i,t}, \tag{B.11}$$

where Equation (B.10) is the model's measurement equation and Equation (B.11) is the model's state equation; $\eta_{i,t}$ is a zero-mean independently and identically normally distributed variable

[7] We generate 1000 such posterior draws.

with variance $\sigma_{i,\eta}$; and $\mu_{i,t}$ is a zero-mean independently and identically normally distributed variable with variance $\sigma_{i,\epsilon}$ which represents the DGP for the unobserved $\epsilon_{i,t}$. The time-varying dimension of this state-space model expresses itself through the time-varying nature of $\zeta_i LOA_{t-1}$ which can be viewed as a time-varying coefficient on $\epsilon_{i,t}$.

We estimate state-space Model (B.10)-(B.11) by applying the Kalman filter to this model to find the values for ζ_i, $\sigma_{i,\eta}$, and $\sigma_{i,\mu}$ that maximize the likelihood function for $\epsilon_i \equiv \{\epsilon_{i,2}, ..., \epsilon_{i,T}\}$ ($\mathbb{P}(\epsilon_i \mid \zeta_i, \sigma_{i,\eta}, \sigma_{i,\mu}, LOA_i, \hat{v}_i)$ where $LOA_i \equiv \{LOA_{i,1}, ..., LOA_{i,T-1}\}$ and $\hat{v}_i \equiv \{\hat{v}_{i,2}, ..., \hat{v}_{i,T}\}$). Our interest lies in the MSE-optimal smoothed $\epsilon_{i,t}$ estimate, which we denote by $\hat{\epsilon}_{i,t}$, obtained from applying this Kalman filter estimation procedure. We now turn to how this estimation is done in our Bayesian setting.

Treatment of Hyperparamters ζ_i, $\sigma_{i,\eta}$, and $\sigma_{i,\mu}$. To simulate posterior draws of ϵ_i, we need to simulate posterior draws from the joint posterior probability distribution of ϵ_i, ζ_i, $\sigma_{i,\eta}$, and $\sigma_{i,\mu}$, $\mathbb{P}(\epsilon_i, \zeta_i, \sigma_{i,\eta}, \sigma_{i,\mu} \mid LOA_i, \hat{v}_i)$, and then use the Kalman filter smoother to obtain posterior draws of ϵ_i. Bayes' law dictates that $\mathbb{P}(\epsilon_i, \zeta_i, \sigma_{i,\eta}, \sigma_{i,\mu} \mid LOA_i, \hat{v}_i) = \mathbb{P}(\zeta_i, \sigma_{i,\eta}, \sigma_{i,\mu} \mid \epsilon_i, LOA_i, \hat{v}_i)\mathbb{P}(\epsilon_i \mid LOA_i, \hat{v}_i)$ and that $\mathbb{P}(\zeta_i, \sigma_{i,\eta}, \sigma_{i,\mu} \mid \epsilon_i, LOA_i, \hat{v}_i) \propto \mathbb{P}(\epsilon_i \mid \zeta_i, \sigma_{i,\eta}, \sigma_{i,\mu}, LOA_i, \hat{v}_i)\mathbb{P}(\zeta_i, \sigma_{i,\eta}, \sigma_{i,\mu} \mid LOA_i, \hat{v}_i)$.

Since we have knowledge of $\mathbb{P}(\epsilon_i \mid LOA_i, \hat{v}_i)$ from the previous section's estimation, all we need in order simulate posterior draws from $\mathbb{P}(\epsilon_i, \zeta_i, \sigma_{i,\eta}, \sigma_{i,\mu} \mid LOA_i, \hat{v}_i)$ is to know $\mathbb{P}(\zeta_i, \sigma_{i,\eta}, \sigma_{i,\mu} \mid \epsilon_i, LOA_i, \hat{v}_i)$. Following the approach of Giannone et al. (2015) and Miranda-Agrippino and Ricco (2021), we treat ζ_i, $\sigma_{i,\eta}$, and $\sigma_{i,\mu}$ as additional model parameters for which we specify a trivariate uniform prior probability distribution and estimate them via the Kalman filter as the maximizers of the posterior likelihood $\mathbb{P}(\zeta_i, \sigma_{i,\eta}, \sigma_{i,\mu} \mid \epsilon_i, LOA_i, \hat{v}_i)$, in the spirit of hierarchical modeling. Our assumed flat prior for the joint prior distribution of ζ_i, $\sigma_{i,\eta}$, and $\sigma_{i,\mu}$ (conditional on ϵ_i,LOA_i, and $\hat{v}_i$), $\mathbb{P}(\zeta_i, \sigma_{i,\eta}, \sigma_{i,\mu} \mid \epsilon_i, LOA_i, \hat{v}_i)$, implies equivalency between maximizing $\mathbb{P}(\epsilon_i \mid \zeta_i, \sigma_{i,\eta}, \sigma_{i,\mu}, LOA_i, \hat{v}_i)$ and $\mathbb{P}(\zeta_i, \sigma_{i,\eta}, \sigma_{i,\mu} \mid \epsilon_i, LOA_i, \hat{v}_i)$ (with respect to ζ_i, $\sigma_{i,\eta}$, and $\sigma_{i,\mu}$).

Estimation of $\mathbb{P}(\epsilon_i \mid LOA_i, \hat{v}_i)$. We are now in position to describe the estimation of the posterior distribution of the smoothed idiosyncratic demand shocks ϵ_i. As noted on Page 16, this posterior distribution can be obtained from $\mathbb{P}(\epsilon_i, \zeta_i, \sigma_{i,\eta}, \sigma_{i,\mu} \mid LOA_i, \hat{v}_i)$ by using the Kalman filter

smoother to produce ϵ_i from the posterior draws of $\hat{v}_i$, ζ_i, $\sigma_{i,\eta}$, and $\sigma_{i,\mu}$.

In particular, for each posterior draw of B_i, we compute $\hat{v}_i$ from Equation (B.5) and perform an unconstrained Kalman filter estimation of Model (B.10)-(B.11) which provides us with estimates of ζ_i, $\sigma_{i,\eta}$, and $\sigma_{i,\mu}$ that we denote by $\zeta_{i,MLE}$, $\sigma_{i,\eta,MLE}$, and $\sigma_{i,\mu,MLE}$, respectively. This Kalman filter estimation provides us with MLE estimates of ζ_i, $\sigma_{i,\eta}$, and $\sigma_{i,\mu}$ which we then feed into the Kalman filter smoother to produce an estimated series of $\epsilon_{i,t}$s (denoted by $\hat{\epsilon}_{i,t}$s) which are then used to construct the posterior distribution of GIV shocks as $\varrho_{GIV,t} = \sum_{i=1}^{13} \hat{\epsilon}_{i,t} w_i - \sum_{i=1}^{13} \hat{\epsilon}_{i,t} \frac{1}{13}$ (normalized to have unit standard deviation), where w_i is II i's share of swap flows' average volume in the sum of IIs' average volumes. We now turn to the estimation of the LOA-dependent effects of this GIV shock from Equation inserted into the estimation of Equation (B.4).

Posterior Simulator. In light of the above-described prior formulation and associated estimation procedure, the posterior simulator for $\varrho_{GIV,t}$ can be described as follows:

1. Do Steps 1-3 from the posterior simulator of Equation (B.1) and obtain B_i and the resulting residual $\hat{v}_t$.

2. Perform an unconstrained Kalman filter estimation of Model (B.10)-(B.11) and obtain the Kalman-filter-smoothed estimate $\hat{\epsilon}_{i,t}$.

3. Construct the GIV shock as $\varrho_{GIV,t} = \sum_{i=1}^{13} \hat{\epsilon}_{i,t} w_i - \sum_{i=1}^{13} \hat{\epsilon}_{i,t} \frac{1}{13}$.

4. Repeat steps 1-3 a large number of times and collect the drawn $\varrho_{GIV,t}$s.[8]

B.5 Equation (B.4)

Companion Form of Specification. Let the set of the parameters (coefficient vector and residual standard deviation) to be estimated from Equation (B.4) be given by Q_h and $\sigma_{u,h}$. Equation (B.4)

[8] We generate 1000 such posterior draws.

can then be written in companion form as follows:[9]

$$Y_h = X_h Q_h + \zeta_h, \tag{B.12}$$

where h is the regression's rolling horizon with $h = 1, ..., 10$; $Y_h = [b_{h-1} - b_{-1}, b_h - b_0, ..., b_T - b_{T-h}]'$; $X_h = [X_1, ..., X_{T-h+1}]'$, with $X_t = [1, ..., LOA_{t-1}]'$; $Q_h = [\alpha_h, ..., \beta_h]'$; and $\zeta_h = [u_h, ..., u_T]'$. Q_h here represents the coefficient vector of Equation (B.4) and $\sigma^2_{u,h}$ is the variance of u_{t+h} (the residual from this equation).

Specification of Uninformative Prior. We assume the following normal-inverse Wishart prior distribution for these parameters:

$$vec(Q_h) \mid \sigma^2_{u,h} \sim N(vec(\bar{Q}_{0,h}), \sigma^2_{u,h} \times N_0^{-1}), \tag{B.13}$$

$$\sigma^2_{u,h} \sim IW_1(m_0 S_{0,h}, m_0), \tag{B.14}$$

where N_0 is a 4 x 4 positive definite matrix (4 being the number of coefficients in Equation (B.4)); S_0 is a variance scalar; and $m_o > 0$. As shown by Uhlig (1994), the latter prior implies the following posterior distribution:

$$vec(Q_h) \mid \sigma^2_{u,h} \sim N(vec(\bar{Q}_h), \sigma^2_{u,h} \times N_h^{-1}), \tag{B.15}$$

$$\sigma^2_{u,h} \sim IW_1(m_h S_h, m_h), \tag{B.16}$$

where $m_h = T - h + 1 + m_0$; $N_h = N_0 + X_h' X_h$; $\bar{Q}_h = N_h^{-1}(N_0 \bar{Q}_{0,h} + X_h' X_h \hat{Q}_h)$; $S_h = \frac{m_0}{m_h} S_{0,h} + \frac{T-h+1}{m_h} \hat{\sigma}^2_{u,h} + \frac{1}{m_h}(\hat{Q}_h - \bar{Q}_{0,h})' N_0 N_h^{-1} X_h' X_h (\hat{Q}_h - \bar{Q}_{0,h})$, where $\hat{Q}_h = (X_h' X_h)^{-1} (X_h)' Y$ and $\hat{\sigma}^2_{u,h} = (Y_h - X_h \hat{Q}_h)'(Y_h - X_h \hat{Q}_h)/(T - h + 1)$.

We use a weak prior, i.e., $m_0 = 0$, $N_0 = 0$, and arbitrary $S_{0,h}$ and $\bar{Q}_{0,h}$. This implies that the prior distribution is proportional to $\sigma^2_{u,h}$ and that $m_h = I \times (T - h + 1)$, $S_h = \hat{\sigma}^2_{u,h}$, $\bar{Q}_h = \hat{Q}_h$, and $N_h = X_h' X_h$. Due to the temporal correlations of the error term u_{t+h}, the likelihood function is misspecified which in turn requires that the residual variance estimate $\hat{\sigma}^2_{u,h}$ be appropriately

[9] While there is, technically speaking, notational abuse of capital letters Y and X in Equations (B.5) and (B.12), what matters is that they are of course defined and used in the context of their corresponding equations and hence their specific contexts avoid expositional confusion. (The context for the first equation is a single equation while for the second it is a rolling regression context (hence the use of subscript h for Y and X in the second equation).)

modified so as to improve estimation precision (Müller (2013)). Toward this end, we apply a Newey-West correction to $\hat{\sigma}^2_{u,h}$ which accounts for arbitrary temporal correlation of the error term and denote the corrected variance estimate by $\hat{\sigma}^2_{u,h,hac}$.

Posterior Simulator. Given the above-described prior formulation and the correction to $\hat{\sigma}^2_{u,h}$, we are now in position to lay out the posterior simulator for Q_h and $\sigma^2_{u,h}$, which accounts for uncertainty in the estimation of all of the model's equations and can be described as follows:

1. Do Steps 1-3 and 1-4 from the posterior simulators of Equations (B.1) and (B.10)-(B.11), respectively, and obtain $\varrho_{GIV,t}$ (whose standardized value is to be used as explanatory variables for the next two steps).

2. Draw $\sigma^2_{u,h}$ from an $IW_1((T-h+1)\hat{\sigma}^2_{u,h,hac}, (T-h+1))$ distribution.

3. Draw Q_h from the conditional distribution $MN(\hat{Q}_h, \sigma^2_{u,h} \times (X_h' X_h)^{-1})$.

4. Repeat Steps 1-3 a large number of times and collect the drawn Q_h's and $\sigma^2_{u,h}$'s.[10]

B.6 Smoothing of Impulse Responses

The high-frequency nature of our data combined with using a local projection estimation approach produce rather jagged raw impulse responses. It is therefore important and warranted to smooth the raw impulse responses with a data-dependent smoothing procedure that suitably integrates into our Bayesian framework. We now turn to present this procedure.

General Setting and Objective. Let $\Theta_L = [\Xi_{L,1}, \Xi_{L,2}, ..., \Xi_{L,10}]'$ and $\Theta_I = [\Xi_{I,1}, \Xi_{I,2}, ..., \Xi_{I,10}]'$ denote the raw linear and non-linear (interaction term based) impulse response 10×1 vectors, respectively. The posterior simulator of Steps 1-4 shows how to obtain posterior draws for these vectors and hence effectively gives knowledge of its posterior probability distribution $\mathbb{P}(\Theta_j \mid data)$ $(j = [L, I])$. We assume the following smooth trend, state-space model for $\Xi_{j,h}$ (with h representing

[10] The number of posterior draws is 1000, as this posterior simulator generates a posterior draw for each of the 1000 drawn GIV shocks.

the horizon and $j = [L, I]$):

$$\Xi_{j,h} = \tilde{\Xi}_{j,h} + \tau_{j,h}, \tag{B.17}$$

$$\tilde{\Xi}_{j,h} = 2\tilde{\Xi}_{j,h-1} - \tilde{\Xi}_{j,h-2} + \gamma_{j,h}, \tag{B.18}$$

where Equation (B.17) is the model's measurement equation and Equation (B.18) is the model's state equation; $\tilde{\Xi}_{j,h}$ is the smoothed impulse response at horizon h whose first-difference follows a random walk with shock $\gamma_{j,h}$, which is a zero-mean independently and identically normally distributed variable with variance σ_γ; and $\tau_{j,h}$ is a zero-mean independently and identically normally distributed variable with variance σ_τ which represents the noise embodied in the raw impulse response function (IRF). Letting $\tilde{\Theta}_j = [\tilde{\Xi}_{j,1}, \tilde{\Xi}_{j,2}, ..., \tilde{\Xi}_{j,10}]'$ denote the smoothed impulse response 10×1 vector, our main interest and objective can be described as lying in simulating $\tilde{\Theta}_j$ from its posterior probability distribution $\mathbb{P}(\tilde{\Theta}_j \mid data)$.

Inherently, the issue of jagged responses in local projection estimations arises for the non-impact, rolling horizons and it usually exacerbates as the rolling horizon gets longer owing to the increased estimation variance from the overlapping nature of rolling regressions such as Equation (B.4). However, impact response estimation does not suffer from this issue given that it corresponds to a standard, non-overlapping regression specification. Hence, it is important to use a smoothing procedure that maintains the raw impact response while smoothing the subsequent horizon responses.[11] Toward this end, we formulate state space Model (B.17)-(B.18) in a time-varying manner such that for $h = 1$ $\sigma_\tau = 0$ σ_γ while for $h = 2, ..., 10$ we leave σ_τ σ_γ as unrestricted parameters. The time-varying nature of this formulation is manifested through the varying nature of the model's parameters of interest and the Kalman filter estimation procedure accounts for this thereby resulting in an estimated impulse response vector where its first element equals the raw impact estimated response and the subsequent elements are smoothed estimated responses.[12]

Treatment of Hyperparamters σ_τ and σ_γ. To simulate posterior draws of $\tilde{\Theta}_j$, we need to simulate posterior draws from the joint posterior probability distribution of Θ_j, σ_τ, and σ_γ, $\mathbb{P}(\Theta_j, \sigma_\tau, \sigma_\gamma \mid$

[11] Any otherwise developed smoothing procedure would potentially generate a smoothed impact effect that can meaningfully deviate from the raw one.

[12] Note that the Bayesian raw impact estimated response would only equal its baseline, classical counterpart from the text in population. In finite samples some deviation between them is to be expected.

$data$), and then use the Kalman filter smoother to obtain posterior draws of $\tilde{\Theta}_j$. Bayes' law dictates that $\mathbb{P}(\Theta_j, \sigma_\tau, \sigma_\gamma \mid data) = \mathbb{P}(\sigma_\tau, \sigma_\gamma \mid \Theta_j, data)\mathbb{P}(\Theta_j \mid data)$ and that $\mathbb{P}(\sigma_\tau, \sigma_\gamma \mid \Theta_j, data) \propto \mathbb{P}(\Theta_j \mid \sigma_\tau, \sigma_\gamma, data)\mathbb{P}(\sigma_\tau, \sigma_\gamma \mid data)$.

Since we have knowledge of $\mathbb{P}(\Theta_j \mid data)$ from the previous section's estimation, all we need in order simulate posterior draws from $\mathbb{P}(\Theta_j, \sigma_\tau, \sigma_\gamma \mid data)$ is to know $\mathbb{P}(\sigma_\tau, \sigma_\gamma \mid \Theta_j, data)$. Following the approach of Giannone et al. (2015) and Miranda-Agrippino and Ricco (2021), we treat σ_τ and σ_γ as additional model parameters for which we specify a bivariate uniform prior probability distribution and estimate them via the Kalman filter as the maximizers of the posterior likelihood $\mathbb{P}(\sigma_\tau, \sigma_\gamma \mid \Theta_j, data)$, in the spirit of hierarchical modeling. Our assumed flat prior for the joint prior distribution of σ_τ and σ_γ (conditional on the observed data), $\mathbb{P}(\sigma_\tau, \sigma_\gamma \mid data)$, implies equivalency between maximizing $\mathbb{P}(\Theta_j \mid \sigma_\tau, \sigma_\gamma, data)$ and $\mathbb{P}(\sigma_\tau, \sigma_\gamma \mid \Theta_j, data)$ (with respect to σ_τ and σ_γ).

Estimation of $\mathbb{P}(\tilde{\Theta}_j \mid data)$. We are now in position to describe the estimation of the posterior distribution of the smoothed IRF $\tilde{\Theta}_j$. As noted on Page 16, this posterior distribution can be obtained from $\mathbb{P}(\Theta_j, \sigma_\tau, \sigma_\gamma \mid data)$ by using the Kalman filter smoother to produce $\tilde{\Theta}_j$ from the posterior draws of Θ_j, σ_τ, and σ_γ.

In particular, for each posterior draw of $\Xi_{j,h}$, we perform an unconstrained Kalman filter estimation of Model (B.17)-(B.18) which provides us with estimates of σ_τ and σ_γ that we denote by $\sigma_{\tau,MLE}$ and $\sigma_{\gamma,MLE}$, respectively. This Kalman filter estimation provides us with estimates of σ_τ and σ_γ which we then feed into the Kalman filter smoother to produce a smoothed IRF. Applying this smoothing procedure to the 1000 posterior draws of raw IRFs results in the sought after posterior distribution of smoothed IRFs.

FEV Estimation Method. For the forecast error variance (FEV) decomposition estimation, we utilize the estimated (smoothed) LOA-dependent impulse responses to compute the LOA-dependent FEV contributions of our swap demand shock as follows:

$$\mathbb{C}_{LOA,h} = \frac{\left(\hat{\tilde{\Xi}}_{L,0} + 2\hat{\tilde{\Xi}}_{I,0}\right)^2 + ... + \left(\hat{\tilde{\Xi}}_{L,h} + 2\hat{\tilde{\Xi}}_{I,h}\right)^2}{\mathbb{V}(b_{t+h} - b_{t-1} \mid LOA)}, \tag{B.19}$$

$$\mathbb{C}_{NLOA,h} = \frac{\left(\hat{\tilde{\Xi}}_{L,0} - 1.4\hat{\tilde{\Xi}}_{I,0}\right)^2 + ... + \left(\hat{\tilde{\Xi}}_{L,h} - 1.4\hat{\tilde{\Xi}}_{I,h}\right)^2}{\mathbb{V}(b_{t+h} - b_{t-1} \mid NLOA)}, \tag{B.20}$$

where $\hat{\Xi}_{L,h}$ and $\hat{\Xi}_{I,h}$ are the estimated linear and nonlinear (interaction-term) impulse response coefficients from Equation (B.4); LOA and $NLOA$ correspond to the LOA and no LOA states, respectively; and $\mathbb{V}(b_{t+h} - b_{t-1} \mid LOA)$ and $\mathbb{V}(b_{t+h} - b_{t-1} \mid NLOA)$ represent the variances of IIs' cross-currency basis' accumulated differences conditional on the LOA and no LOA states.

Operationally, we define the LOA state for the FEV estimation as the group of observations where the LOA series values are above or equal to the LOA series's 92.8th percentile. The rationale for this definition is based on the fact that the LOA state for the impulse response estimation is defined by the LOA variable being equal to its 96.4th percentile value. (The LOA variable's 2 standard deviation value corresponds to its 96.4th percentile.) Hence, we define the variance conditional on this state as the variance that results from considering observations that closely and symmetrically surround the impulse response estimation's LOA state value but at the same time delivers a sufficient number of observations for FEV estimation. The no LOA state is symmetrically defined for the FEV estimation as the group of observations where the LOA series values are below or equal to the LOA series's 7.2th percentile.

Results. Figures B.1 and B.2 present LOA-dependent impulse responses and FEVs of IIs' basis with respect to an aggregate FX swap demand shock (as measured by our GIV shock). We scale all of the impulse responses in our empirical analysis such that the impact linear response of the IIs' aggregate FX swap flow variable is equal to one standard deviation of this variable (542.5 millions dollars).

The results indicate a short-lived basis response, with the basis widening significantly for only 3 trading days. And the contribution of the GIV shock to the variation of the basis is low on impact, standing at 3.3%, and peaks at a slightly higher share of 5.8% after 7 days. We do not view these low shares as evidence going against our paper's message given that our claim is not that the LOA-dependent FX swap demand channel explains the bulk of the variation in the basis but rather that our data allows us to meaningfully uncover the presence of this channel in the data.

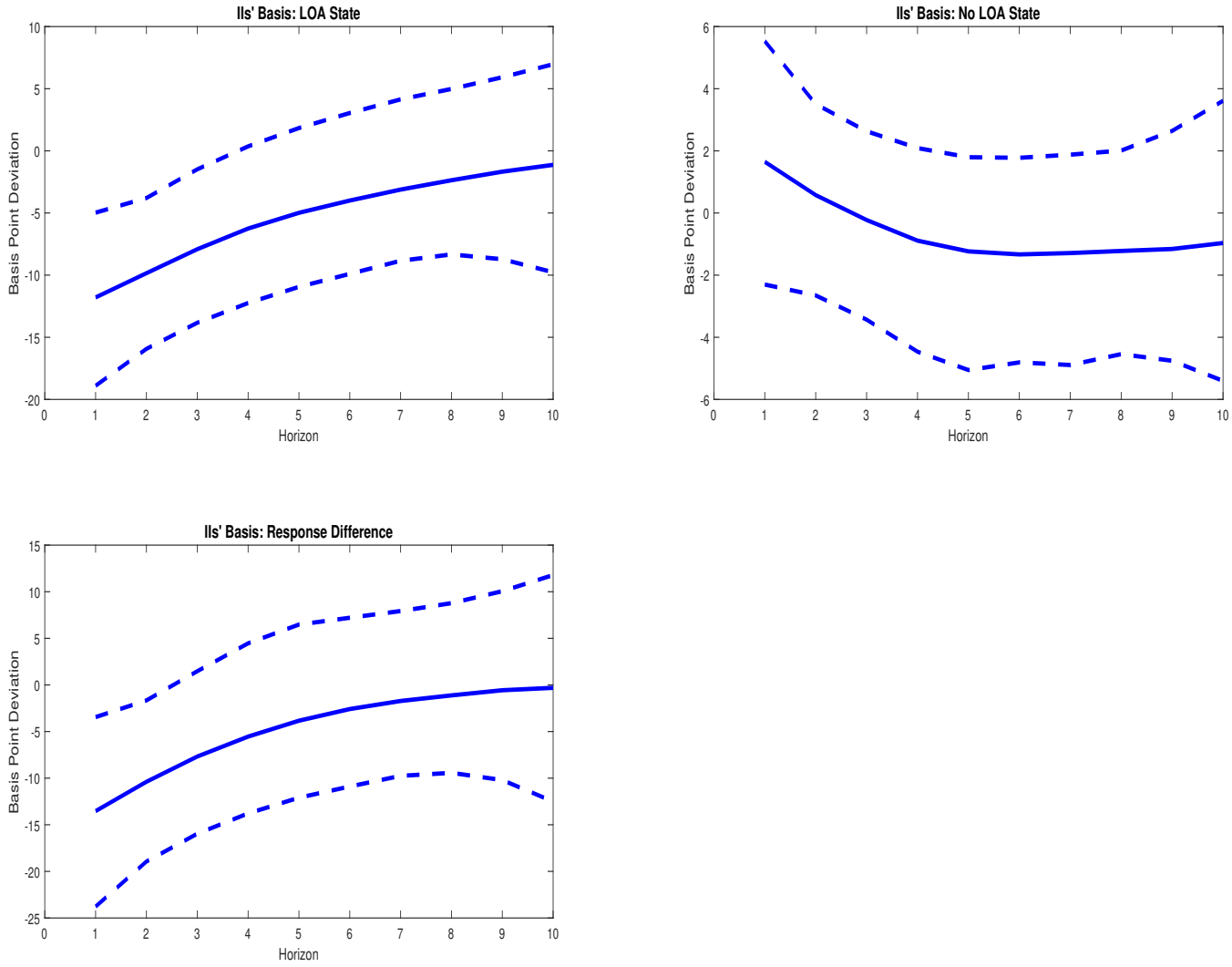
What can explain the temporary basis response from Figure B.1? To properly answer this question, the basis response must be viewed together with the swap flow LOA-dependent response. Toward this end, the LOA-dependent response of IIs' aggregate open FX swap position is shown

in Figure B.3. This response is obtained by replacing the dependent variable $b_{t+h-1} - b_{t-1}$ from Equation (B.4) with $SP_{t+h-1} - SP_{t-1}$ (IIs' aggregate open FX swap position) and also adding a time trend to the RHS of this equation.

The persistence of IIs' open swap position's response in the LOA state is quite low, with the swap position dropping to 72.5 millions dollars at the 4th horizon from an impact value of 188.5 million dollars. And the response is no longer significant at the 7th and 8th horizons, regains significance at the 9th horizon, and then returns to being insignificant. The low persistence of the position's response provides further support for the interpretation of IIs' buying of swap-linked dollars as a temporary action which essentially takes place entirely on impact, after which the associated open position is rather quickly eliminated. And this interpretation is consistent with the largely expedited closing of the basis from Figure B.1, i.e., the low persistence of the GIV shock in the LOA state can serve as one sensible explanation for the rather temporary basis widening in this state.[13]

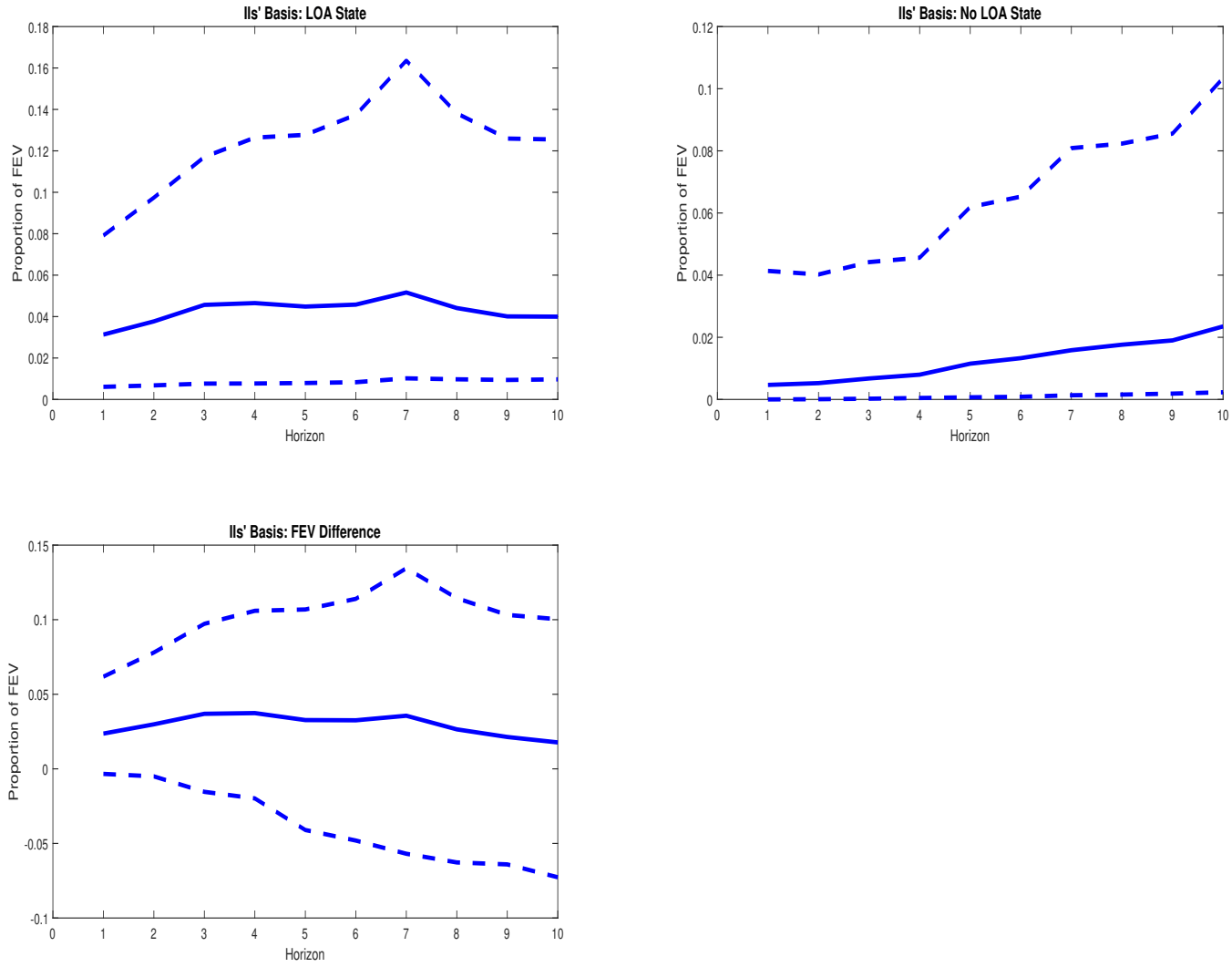
[13]Note that the swap position response in the no LOA (linear) state does not posses this low persistence, remaining very significant for all considered horizons. The apparent differential persistence of the GIV shock in the two states speaks to the importance of focusing on the impact effects of the GIV shock for the identification of the LOA-dependent FX swap demand channel, as we do in the paper, as in the impact period the demand shock faces a differentially sloped FX swap curve while in later periods factors such as differential persistence (as well as supply slope variation coming from arbitragers looking to seize arbitrage opportunities present in the LOA state) come into play and sully the identification of the channel.

Figure B.1: **LOA-Dependent Impulse Responses of IIs' Aggregate Cross-Currency Basis to an Aggregate FX Swap Demand Shock**.



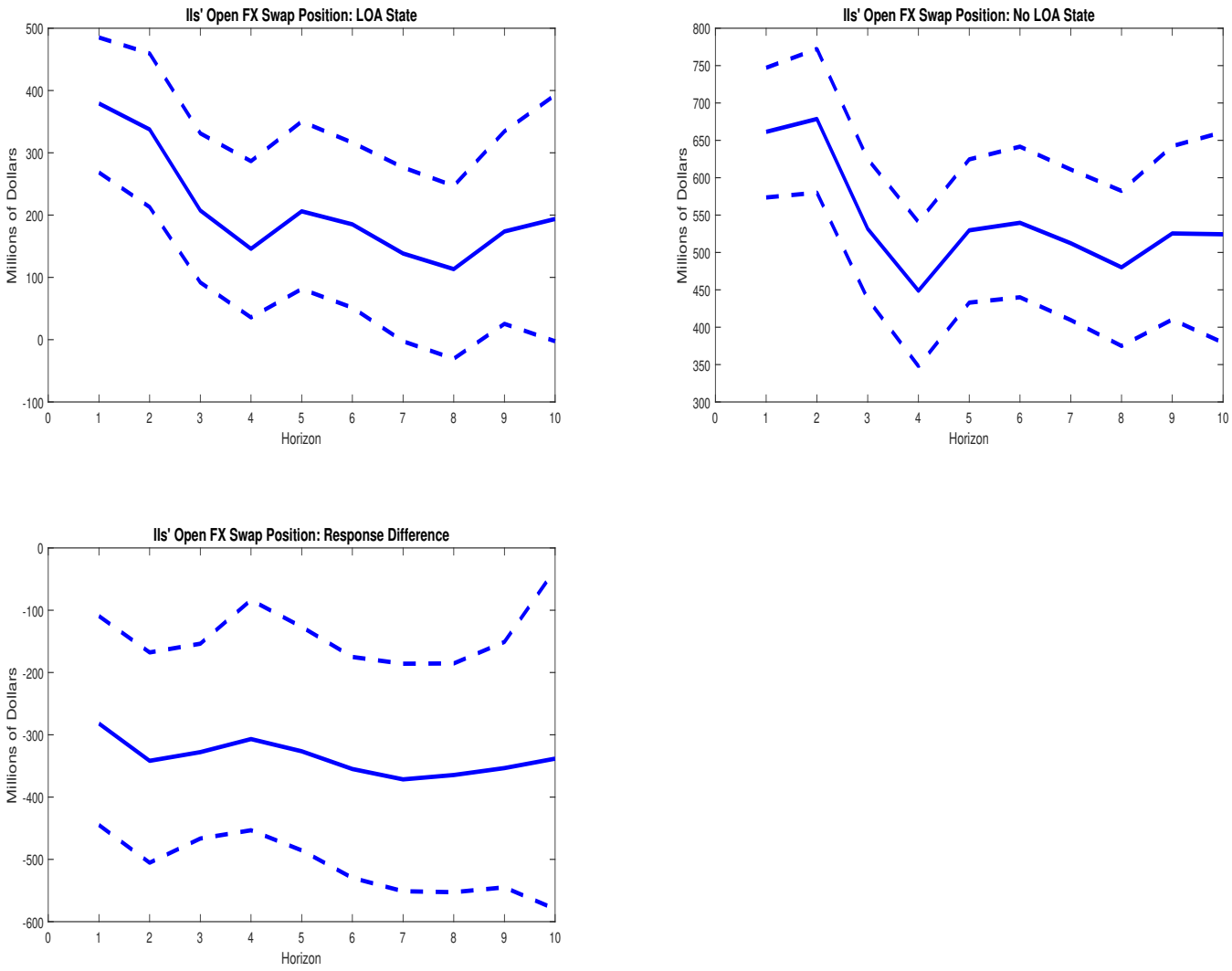
Notes: This figure presents the LOA-dependent impulse responses of IIs' aggregate cross currency basis to the aggregate FX swap demand shock from the model described by Equations (B.1)-(B.4). Solid lines represent median values of posterior distribution of impulse responses and dashed lines depict the 95% bands of this distribution. The first and second columns show the responses in the LOA and no LOA states, respectively; and the third column shows the response differences across these two states. Responses are in terms of deviations from pre-shock values (basis point deviations). Horizon (on x-axis) is in days.

Figure B.2: **LOA-Dependent FEV Shares of IIs' Aggregate Cross-Currency Basis Attributable to the Aggregate FX Swap Demand Shock**.



Notes: This figure presents the FEV share of IIs' aggregate cross-currency basis that is attributable to the aggregate FX swap demand shock from the model described by Equations (B.1)-(B.4). Solid lines represent median values of posterior distribution of FEVs and dashed lines depict the 95% bands of this distribution. The first and second columns show the FEV contributions in the LOA and no LOA states, respectively; and the third column shows the FEV contribution differences across these two states. Horizon (on the x-axis) is in days and the FEV share is on the y-axis.

Figure B.3: **LOA-Dependent Impulse Responses of IIs' Aggregate Open FX Swap Position to an Aggregate FX Swap Demand Shock**.



Notes: This figure presents the LOA-dependent impulse responses of IIs' aggregate open swap position to the aggregate FX swap demand shock from the model described by Equations (B.1)-(B.4) where the outcome variable in the latter equation (accumulated difference in IIs' basis) is now replaced by the accumulated difference in IIs' open FX swap position (i.e., $SP_{t+h-1} - SP_{t-1}$) and a time trend is also added to the RHS of this equation. Solid lines represent median values of posterior distribution of impulse responses and dashed lines depict the 95% bands of this distribution. The first and second columns show the responses in the LOA and no LOA states, respectively; and the third column shows the response differences across these two cases. Responses are in terms of deviations from pre-shock values (in millions of dollars terms). Horizon (on x-axis) is in days.

Appendix C Coefficient Estimates Results

This appendix presents the estimates of the linear coefficient (Ξ_L), interaction coefficient (Ξ_I), and coefficient on interacting variable LOA_{t-1} (β) from estimation of Equations (8) and (7) from the text (Table C.1); from estimation of Equation (8) from the text where the basis variable in this equation is replaced with each sector's swap flow variable and a time trend is added to the RHS of these equations (Table C.2); from estimation of Equation (8) from the text for various alternative specifications (Table C.3 - see notes for this table for details about these specifications); and from estimation of Equations (9) and (10) from the text (Table C.4).

Table C.1: **Granular-With-Controls Estimation Results: Coefficient Estimates.**

	IIs' Aggregate Basis (in Basis Points)	
Coefficient	GIV-With-Controls	Bartik-With-Controls
Linear Coefficient	-3.7***	-2.8***
	(1.3)	(0.7)
Interaction Coefficient	-4.2**	-2.4**
	(1.8)	(1.0)
LOA_{t-1} Coefficient	-0.2	-0.1
	(0.2)	(0.2)
R^2	1.8%	1.2%
Obs	2,648	2,648
	IIs' Aggregate Swap Flows (in Millions of Dollars)	
Coefficient	GIV-With-Controls	Bartik-With-Controls
Linear Coefficient	542.5***	542.5***
	(21.3)	(12.2)
Interaction Coefficient	-101.6***	-119.3***
	(23.1)	(14.5)
LOA_{t-1} Coefficient	6.0	2.4
	(4.5)	(5.8)
R^2	37.3%	55.9%
Obs	2,648	2,648

Notes: The first panel of this table presents the estimates of the linear coefficient (Ξ_L), interaction coefficient (Ξ_I), and coefficient on interacting variable LOA_{t-1} (β) from estimation of Equations (8) and (7) from the text, respectively, where the two shocks are the difference between the size-weighted- and equal-weighted-average (GIV) and the equal-weighted-average (Bartik) of the idiosyncratic II-level demand shocks estimated from Equations (4)-(6) from the text. The second panel of the table shows the corresponding estimates obtained by replacing the basis variable in Equations (8) and (7) from the text with IIs' aggregate swap flows one and adding a time trend to the RHS of these equations. As in all of this paper's estimations, the linear and interaction coefficients' estimates are scaled such that the linear demand effect on the IIs' aggregate swap flows variable is equal to one standard deviation of this variable (542.5 million dollars). Numbers in parentheses represent standard errors computed from the heteroskedasticity- and autocorrelation-consistent procedure of Newey and West (1987) with the truncation lag selected from the data-driven procedure from Andrews (1991). *, **, and *** represent significance levels at the 10%, 5%, and 1% levels.

Table C.2: **GIV-With-Controls Estimation Results: Sectoral Swap Flows: Coefficient Estimates.**

Coefficient	IIs	Local Banks	Foreigners	MFs and ETFs	HFs	Real
Linear Coefficient	542.5***	-398.5***	-97.8***	-4.6	-12.7	10.1
	(21.3)	(40.9)	(24.8)	(5.4)	(26.4)	(9.3)
Interaction Coefficient	-101.6***	152.6***	-41.5	-1.9	-27.3	5.9
	(23.1)	(41.1)	(35.7)	(4.1)	(19.5)	(6.1)
LOA_{t-1} Coefficient	6.0	-8.4	1.5	-1.0	-4.9	-0.3
	(4.5)	(11.9)	(11.1)	(1.3)	(3.2)	(1.3)
R^2	37.3%	4.9%	0.8%	0.1%	0.1%	0.1%
Obs	2,648	2,648	2,648	2,648	2,648	2,648

Notes: This table presents the estimates of the linear coefficient (Ξ_L), interaction coefficient (Ξ_I), and coefficient on interacting variable LOA_{t-1} (β) from estimation of Equation (8) where the basis variable in this equation is replaced with each sector's swap flow variable and a time trend is added to the RHS of these equations. The two shocks are the difference between the size-weighted- and equal-weighted-average (GIV) and the equal-weighted-average (Bartik) of the idiosyncratic II-level demand shocks estimated from Equations (4)-(6) from the text. For completeness, IIs' swap flows' responses are presented in the first column of the table, followed by the corresponding responses for local banks (second column); foreigners (third column); mutual funds (MFs) and exchange trade funds (ETFs) (abbreviated by MFs in the fourth column); hedge funds and proprietary trading firms (abbreviated by HFs in the fifth column); and the real sector (seventh column). Numbers in parentheses represent standard errors computed from the heteroskedasticity- and autocorrelation-consistent procedure of Newey and West (1987) with the truncation lag selected from the data-driven procedure from Andrews (1991). *, **, and *** represent significance levels at the 10%, 5%, and 1% levels.

Table C.3: **GIV-With-Controls Estimation Results: Robustness Checks: Coefficient Estimates.**

	IIs' Aggregate Basis (in Basis Points)				
Coefficient	Alternative LOA	Post-GFC	Pre-COVID	Shorter Lags	Longer Lags
Linear Coefficient	-3.4***	-2.6***	-3.3***	-3.4**	-3.8***
	(1.2)	(0.8)	(0.8)	(1.3)	(1.3)
Interaction Coefficient	-3.3***	-3.3***	-2.7**	-3.7**	-4.0**
	(1.2)	(1.2)	(1.3)	(1.8)	(1.7)
LOA_{t-1} Coefficient	-0.2	0.0	-0.2	-0.2	-0.2
	(0.2)	(0.2)	(0.2)	(0.2)	(0.2)
R^2	1.3%	1.4%	0.9%	1.7%	1.7%
Obs	2,648	2,392	2,204	2,652	2,640
	IIs' Aggregate Swap Flows (in Millions of Dollars)				
Coefficient	Alternative LOA	Post-GFC	Pre-COVID	Shorter Lags	Longer Lags
Linear Coefficient	542.5***	542.5***	542.5***	542.5***	542.5***
	(14.7)	(19.4)	(22.0)	(20.9)	(20.3)
Interaction Coefficient	-124.0***	-94.8***	-123.2***	-104.0***	-99.6***
	(18.6)	(20.5)	(23.9)	(24.2)	(22.4)
LOA_{t-1} Coefficient	5.5	6.2	-3.4	6.4	6.6
	(8.0)	(5.1)	(4.2)	(4.0)	(5.7)
R^2	39.2%	34.8%	33.9%	42.2%	35.2%
Obs	2,648	2,392	2,204	2,652	2,640

Notes: The first panel of this table presents the estimates of the linear coefficient (Ξ_L), interaction coefficient (Ξ_I), and coefficient on interacting variable LOA_{t-1} (β) from estimation of Equation (8) from the text for various alternative specifications. The first uses an alternative LOA measure taken from He et al. (2017); the second excludes the GFC period by beginning the sample in 2010; the third excludes the COVID period by truncating the sample at the end of February of 2020; and the fourth and fifth halve and double the number of lags from Equation (3) from the text, respectively. The second panel of the table shows the corresponding responses of IIs' aggregate swap flows obtained by replacing the basis variable in Equation (8) from the text with the swap flow one and adding a time trend to the RHS of these equations. As in all of this paper's estimations, the linear and interaction coefficients' estimates are scaled such that the linear demand effect on the IIs' aggregate swap flows variable is equal to one standard deviation of this variable (542.5 million dollars). Numbers in parentheses represent standard errors computed from the heteroskedasticity- and autocorrelation-consistent procedure of Newey and West (1987) with the truncation lag selected from the data-driven procedure from Andrews (1991). *, **, and *** represent significance levels at the 10%, 5%, and 1% levels.

Table C.4: **Seasonal Demand Shifter Results: Coefficient Estimates.**

	IIs' Aggregate Basis (in Basis Points)	
Coefficient	Seasonal-No-Controls	Seasonal-With-Controls
Linear Coefficient	4.4	4.6
	(2.9)	(3.3)
Interaction Coefficient	8.6**	7.9**
	(3.7)	(3.7)
LOA_{t-1} Coefficient	0.3	-5.1
	(0.2)	(3.4)
R^2	0.2%	37.9%
Obs	2,650	2,650
	IIs' Aggregate Swap Flows (in Millions of Dollars)	
Coefficient	Seasonal-No-Controls	Seasonal-With-Controls
Linear Coefficient	542.5***	542.5***
	(93.5)	(124.9)
Interaction Coefficient	-224.1***	-150.6*
	(78.4)	(76.9)
LOA_{t-1} Coefficient	-2.6	55.6
	(11.6)	(59.3)
R^2	1.7%	37.8%
Obs	2,650	2,650

Notes: The first panel of this table presents the estimates of the linear coefficient (Ξ_L), interaction coefficient (Ξ_I), and coefficient on interacting variable LOA_{t-1} (β) from estimation of Equations (9) and (10) from the text, respectively. The second panel of the table shows the corresponding estimates obtained by replacing the basis variable in Equations (9) and (10) from the text with IIs' aggregate swap flows one and adding a time trend to the RHS of these equations. As in all of this paper's estimations, the linear and interaction coefficients' estimates are scaled such that the linear demand effect on the IIs' aggregate swap flows variable is equal to one standard deviation of this variable (542.5 million dollars). Numbers in parentheses represent standard errors computed from the heteroskedasticity- and autocorrelation-consistent procedure of Newey and West (1987) with the truncation lag selected from the data-driven procedure from Andrews (1991). *, **, and *** represent significance levels at the 10%, 5%, and 1% levels.

References

Andrews, D. W. K.: 1991, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica* **59**(3), 817–858.

Ben Zeev, N.: 2023, The TFP channel of credit supply shocks, *The Review of Economics and Statistics* **105**(2), 425–441.

Del Negro, M. and Schorfheide, F.: 2011, Bayesian Macroeconometrics, *in* G. K. John Geweke and H. V. Dijk (eds), *The Oxford Handbook of Bayesian Econometrics*, Oxford University Press, p. 293–389.

Du, W. and Schreger, J.: 2022, Chapter 4 - CIP deviations, the dollar, and frictions in international capital markets, *in* G. Gopinath, E. Helpman and K. Rogoff (eds), *Handbook of International Economics: International Macroeconomics, Volume 6*, Vol. 6 of *Handbook of International Economics*, Elsevier, pp. 147–197.

Gabaix, X. and Koijen, R. S. J.: 2023, Granular instrumental variables, *Journal of Political Economy (forthcoming)* .

Giannone, D., Lenza, M. and Primiceri, G. E.: 2015, Prior Selection for Vector Autoregressions, *The Review of Economics and Statistics* **97**(2), 436–451.

He, Z., Kelly, B. and Manela, A.: 2017, Intermediary asset pricing: New evidence from many asset classes, *Journal of Financial Economics* **126**(1), 1–35.

Ivashina, V., Scharfstein, D. S. and Stein, J. C.: 2015, Dollar Funding and the Lending Behavior of Global Banks, *The Quarterly Journal of Economics* **130**(3), 1241–1281.

Liao, G. Y. and Zhang, T.: 2020, The Hedging Channel of Exchange Rate Determination, *International Finance Discussion Papers 1283*, Board of Governors of the Federal Reserve System (U.S.).

McLemore, P.: 2019, Do mutual funds have decreasing returns to scale? evidence from fund mergers, *Journal of Financial and Quantitative Analysis* **54**(4), 1683–1711.

Miranda-Agrippino, S. and Ricco, G.: 2021, The transmission of monetary policy shocks, *American Economic Journal: Macroeconomics* **13**(3), 74–107.

Müller, U. K.: 2013, Risk of bayesian inference in misspecified models, and the sandwich covariance matrix, *Econometrica* **81**(5), 1805–1849.

Newey, W. K. and West, K. D.: 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* **55**(3), 703–708.

Plagborg-Møller, M.: 2016, *Essays in macroeconometrics*, Phd thesis, department of economics, harvard university.

Schrimpf, A. and Sushko, V.: 2019, Sizing up global foreign exchange markets, *BIS quarterly review*, Bank for International Settlements.

Shleifer, A. and Vishny, R. W.: 1997, The limits of arbitrage, *The Journal of Finance* **52**(1), 35–55.

Uhlig, H.: 1994, What macroeconomists should know about unit roots: A bayesian perspective, *Econometric Theory* **10**(3/4), 645–671.

Zha, T.: 1999, Block recursion and structural vector autoregressions, *Journal of Econometrics* **90**(2), 291–316.

Zhu, H.: 2008, Capital Regulation and Banks' Financial Decisions, *International Journal of Central Banking* **4**(1), 165–211.