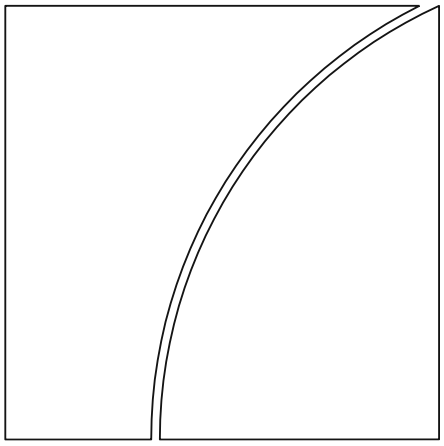

BIS Working Papers

No 1206

Covered interest parity: a forecasting approach to estimate the neutral band

by Juan R. Hernández

Monetary and Economic Department

August 2024

JEL classification: C52, C58, F31, F37, G15, G17

Keywords: Covered interest parity, carry trade, stochastic volatility, predictive distribution

BIS Working Papers are written by members of the Monetary and Economic Department of the Bank for International Settlements, and from time to time by other economists, and are published by the Bank. The papers are on subjects of topical interest and are technical in character. The views expressed in them are those of their authors and not necessarily the views of the BIS.

This publication is available on the BIS website (www.bis.org).

ISSN 1020-0959 (print)
ISSN 1682-7678 (online)

Covered Interest Parity: A Forecasting Approach to Estimate the Neutral Band*

Juan R. Hernández
Centro de Investigación y Docencia Económicas (CIDE);
Technical Adviser, Bank for International Settlements (BIS) Americas Office

August 2024

Abstract

The neutral band is the interval where deviations from covered interest parity (CIP) are not considered profitable arbitrage opportunities. After the great financial crisis, deviations from CIP are no longer short-lived, exposing some limitations of the previous approaches to estimate the neutral band. In this paper, I argue that the one-step-ahead forecast distribution of deviations from CIP, with a time-varying variance component, provides an intuitive estimate of the neutral band. I use data for the Pound Sterling-US Dollar cross from 2000 to 2021, and find that a stochastic volatility model outperforms several alternative models in terms of fit and forecasting capability. The model estimates neutral band that are intuitive and consistent with market dynamics, widening during financial stress periods and consistent with no arbitrage. The results are maintained when I use data from the Mexican Peso-US Dollar cross.

Keywords: Covered interest parity; carry trade; stochastic volatility; predictive distribution.
J.E.L. classification number. C52, C58, F31, F37, G15, G17.
Address for Correspondence: Juan R. Hernández, Assistant Professor. Centro de Investigación y Docencia Económicas. Carretera México-Toluca 3655, Col. Lomas de Santa Fe. Ciudad de México, CP 01210. México
E-mail: juan.hernandez@cide.edu
Acknowledgements: Nicolás Amoroso, Alfonso Cebreros, Raúl Ibarra, Mauricio Olivares and Vladimir Rodríguez provided helpful discussion and comments on previous drafts. Jon Frost's comments were particularly helpful. Comments from referees, participants at the BIS Americas Office seminar series in 2024, CIDE in 2021, ITAM's Seminario Aleatorio in 2019, the 4th Market Microstuctre-Nonlinear Dynamics Workshop in 2019 are gratefully acknowledged. Andrea Miranda, Mariana Ruiz and Ezequiel Piedras provided excellent research assistance. All remaining errors are my own.

*Earlier versions circulated under the title "Covered Interest Parity: A Stochastic Volatility Approach to Estimate the Neutral Band" and were developed while the author was affiliated to Banco de México. Opinions contained in this paper are my own views and not necessarily those of the BIS, or Banco de México.

1 Introduction

Covered interest parity (CIP) is a cornerstone of international finance and a bellwether of market efficiency. It is used in financial markets to price forward contracts and funds obtained on a covered basis, and to guide a carry trade strategy. CIP is used by central banks as a measure of prevailing liquidity conditions in the foreign exchange (FX) market and is embedded in open economy macro-finance models. CIP is a no-arbitrage condition. A deviation that is sufficiently large will create an arbitrage opportunity, allowing FX market participants to make a profit without incurring any risk. But how large? Deviations from CIP only become a profitable trade or a sign of financial stress, if they are outside an interval known as the "neutral band". The band accounts for the costs derived from required margins, transactions fees and risk.

Prior to the Great Financial Crisis (GFC), transaction costs and risks were low, allowing market participants to trade away deviations from CIP. As a result, deviations were short-lived, and CIP was among the most reliable no-arbitrage relations in international finance. Estimates for the neutral band were relatively narrow and constant. However, the events that unfolded after the GFC seem to have had lasting effects on CIP and, thus, on the neutral band. Deviations are now persistent and have been the norm in covered FX markets for a number of currency crosses with the US Dollar (USD). In this paper, I argue that neutral band estimates that are based on a forecast and allow for changes in volatility are consistent with and the covered FX market dynamics across periods.

Theoretically sound neutral band estimates can prove useful to estimate the magnitude of the costs generated by transactions, regulation and risk (Levich, 2017). Financial authorities can use deviations from CIP as a gauge of liquidity conditions in the FX market (Levich, 2017; Du and Schreger, 2022). However, the neutral band can provide guidance on when these deviations are significant enough to prompt action. For example, in economies with a floating exchange rate regime, financial authorities have a special interest on the appropriate functioning of the FX market (Fratzscher et al., 2019). The neutral band, as a liquidity and cost measure, can aid decisions on FX market intervention. It can also be used in the calibration of models where financial intermediaries face constraints on their balance sheet (Du et al., 2022).

Throughout the paper, I use data on the Pound Sterling-US Dollar (GBP-USD) cross-currency basis, as a measure of deviations from CIP, to compare different estimates for the neutral band. The GBP-USD cross-currency basis has consistently shown deviations from zero throughout most of the 2008-2021 period. I use these data to exploit that the GBP-USD FX market is the fourth-highest in transaction volume according to the BIS (2022), which allows the analysis to be abstracted from high costs associated to low liquidity. Deviations from CIP related to the GBP-USD has been the workhorse data to obtain estimates of the neutral band in the literature, providing useful benchmarks (Frenkel, 1973; Frenkel and Levich, 1975, 1977; Peel and Taylor, 2002). Since the model assessment is conditional on the features of the GBP-USD FX market, I also compare estimates of the neutral band for the Mexican Peso-US Dollar cross currency basis. This is a relatively liquid currency with the features of an emerging market economy currency.

There is now a vast body of literature analysing the possible causes for deviations from CIP, particularly in the post-GFC period (Avdjiev et al., 2019; Du et al., 2018; Cerutti et al., 2021; Cendese et al., 2021). Du and Schreger (2022) provide an excellent survey. However, there are no updated estimates of the neutral band (Levich, 2017). Filling this gap is the main contribution of in this paper. The estimates for different periods of the 20th century were obtained using one of two approaches. First, the "counting" approach from Frenkel and Levich (1975) and Frenkel and Levich (1977) involved declaring the upper and lower limits of the neutral band as the thresholds containing 95% of the measured deviations. This approach is very intuitive, specially while deviations from CIP are temporary, since historical data provides a benchmark to deemed an observation out of norm.

However, deviations from CIP in the post-GFC period are far from short-lived. If researchers estimate the neutral band using past values with the counting approach, they would find that arbitrage opportunities are large, persistent and clustered. Justifying this in FX markets seems difficult. Equally important, changes in regulation have resulted in non-negligible liquidity constraints for risk managers (Du et al., 2018; Avdjiev et al., 2019). An additional layer of difficulty is added by the regulatory requirements since liquidity for a given trading session has to be determined beforehand. I argue that,

given the changes in regulation and the persistence of deviations from CIP, the one-step-ahead forecast distribution of a deviation from CIP is a more intuitive candidate for obtaining neutral band estimates.

The second approach to estimating the neutral band is based on the econometric estimation of the 95% confidence intervals around the measured deviations. Branson (1969) is an example using ordinary least squares, and additional estimates based on linear econometric models are surveyed in Levich (2017). Peel and Taylor (2002) propose to use threshold (non-linear) models. Intuitively, data is used to identify for which values of deviations from CIP a different model is necessary to replicate the data. These are known as threshold estimates, which are then labelled as the lower and upper boundaries of the neutral band.

Both the linear and threshold econometric models focus on the conditional mean of CIP and assume a constant variance. However, a time-varying variance might more accurately capture the changes in risk and risk regulations. By explicitly modelling the volatility processes underlying CIP, I can obtain estimates of the neutral band that widen during periods of financial stress and are consistent with time-varying risk and transaction costs. In this paper, I contribute to this literature by synthesising the two approaches and estimating the neutral band as the one-step ahead 95% predictive distribution for deviations from CIP from 2002 to 2021. The results are not conditional on the currency or FX market structure; I obtain similar results when I use data from the Mexican Peso-US Dollar covered market.

The rest of the paper is organised as follows. In Section 2, I provide a brief outline of CIP theory and some definitions. I then use an asset pricing model with liquidity-constraints from Du et al. (2022) to motivate the neutral band as a forecast with time-varying variance and a description of the data. Section 3 contains details on the econometric models. This includes the details of the four models used to estimate the neutral band, all of which are consistent with no-arbitrage. In Section 4 I present the estimates and diagnostics. I find that a stochastic volatility model produces superior neutral band estimates according to the coverage ratio tests and the log-score statistic. The stochastic volatility model is able to replicate the likelihood of short-lived arbitrage opportunities appearing during financial stress events. In Section 5 I use deviations from CIP in the Mexican Peso - US Dollar market and obtain similar results, despite its different risk profile. Finally, Section 6 offers some concluding remarks.

2 CIP, Cross-Currency Basis and the Neutral Band: Theory and Data

This section formally introduces CIP and outlines the transactions necessary to arbitrage away any deviations from it. Specifically, I discuss *round-trip arbitrage*, as defined by Levich (2017) and frequently explained in international finance literature.[1] Next, I employ an intermediary asset pricing model with regulatory constraints from Du et al. (2022) to motivate the neutral band as a determinant of the revisions of investment opportunities. I then introduce the data used in estimating the neutral band.

2.1 CIP

Let us assume momentarily that transaction costs are zero, there is no risk, and there are no limits to the supply of funds for market participants. Following the notation in Du and Schreger (2022) for transactions maturing τ periods ahead, let S_t be the spot exchange rate and let $F_{t,\tau}$ be the forward rate agreed at period t with τ-months ahead maturity. Both are measured as 1 USD per GBP throughout the paper. Also, let $y_{t,\tau}$ and $y^*_{t,\tau}$ be the interest rate on a (zero-coupon) bond with τ-months maturity denominated in USD and GBP, respectively.

Assuming that individuals can always borrow in USD in order to invest in foreign bonds, and that financial frictions and costs are absent, CIP predicts the interest rate in the cash market to be equal

[1] Deardorff (1979) discusses *one-way-arbitrage*, where market participants will need a known amount of foreign currency at a future date, hence, they will be holding one currency before entering arbitrage transactions and a second one after.

to the implied US Dollar "synthetic" rate

$$1 + y_{t,\tau} = \left(1 + y^*_{t,\tau}\right) \frac{S_t}{F_{t,\tau}}. \tag{2.1}$$

Arbitrageurs are investors that exploit deviations in (2.1) in search of a profit in carry trade. For example, if

$$1 + y_{t,\tau} < \left(1 + y^*_{t,\tau}\right) \frac{S_t}{F_{t,\tau}}. \tag{2.2}$$

the following simultaneous trades will yield a risk-less profit. An investor borrows 1 USD for one period at a rate $y_{t,\tau}$, and buys S_t GBP. Simultaneously, the investor lends S_t GBP at an interest rate $y^*_{t,\tau}$ and enters a forward contract promising to deliver $(1 + y^*_{t,\tau})S_t$ GBP after τ months in exchange for $(1+y^*_{t,\tau})S_t/F_{t,\tau}$ USD. After τ months, the investor holds $(1+y^*_{t,\tau})S_t/F_{t,\tau}$ USD and has a commitment to pay back $(1+y_{t,\tau})$. Since (2.2) holds, the investor has made a profit of $(1+y^*_{t,\tau})S_t/F_{t,\tau} - (1+y_{t,\tau})$ USD.

2.2 Cross-Currency Basis and the Neutral Band

Following the convention in the literature, a wedge in CIP is included as an extra term in the right-hand side of (2.1). The wedge is represented by $x_{t,\tau}$ and is an extra premium on the synthetic rate (Du and Schreger, 2022) with all contracts signed at period t

$$1 + y_{t,\tau} = \left(1 + y^*_{t,\tau} + x_{t,\tau}\right) \frac{S_t}{F_{t,\tau}} \tag{2.3}$$

and the corresponding log-form

$$x_{t,\tau} = y_{t,\tau} - \left(y^*_{t,\tau} - \rho_{t,\tau}\right), \tag{2.4}$$

where $x_{t,\tau}$ is the cross-currency basis (deviations from CIP) and $\rho_{t,\tau} = \log(F_{t,\tau}) - \log(S_t)$ is the forward premium.

Risk-less arbitrage opportunities arise when $x_{t,\tau} \neq 0$ as in (2.2) and undertaking these opportunities re-establishes $x_{t,\tau} = 0$. International finance literature has long acknowledged that $x_{t,\tau} = 0$ may not hold exactly if transaction or regulatory-induced costs and risk exist, and if the supply of funds for market participants is limited. The works of Du et al. (2018); Avdjiev et al. (2019) and Du and Schreger (2022) have found that changes in macro-prudential regulation, particularly limits on leverage imposed on global systemically important banks, may explain the observed $x_{t,\tau} \neq 0$ in the post GFC period. Alternatively, Cerutti et al. (2021) argue that additional factors, such as risk appetite and monetary policy in advanced economies, also contribute to explaining deviations from CIP.

But how large should $|x_{t,\tau}|$ be to attract investors? CIP is considered satisfied as long as $x_{t,\tau}$ remains within a *neutral band*, $b_{t,\tau} = [\underline{b}_{t,\tau}, \overline{b}_{t,\tau}]$, where $\underline{b}_{t,\tau}$ and $\overline{b}_{t,\tau}$ represent its lower and upper limits, respectively (Einzig, 1967; Branson, 1969; Frenkel, 1973; Deardorff, 1979; Peel and Taylor, 2002). If the observed deviation from CIP falls within the neutral band (i.e. profits are smaller than some minimum required), arbitrageurs will not enter any carry trade with certainty. In contrast, if the observed deviation from CIP lies outside the neutral band, arbitrageurs will trade to secure risk-less profits, driving $x_{t,\tau}$ back to $b_{t,\tau}$.[2]

The width of $b_{t,\tau}$ is determined by costs, risks inherent to arbitrage transactions and the availability of funds, all of which may vary over time. In periods of financial stress, $b_{t,\tau}$ widens and connections between financial markets weaken (Levich, 2017). Arbitrage implies that there is no clustering of deviations from CIP outside a valid neutral band (Deardorff, 1979). As a result, the width of $b_{t,\tau}$ will vary over time, conditional on recent available information. However, the width can only be estimated from observed data, an endeavour I undertake below.

[2] Peel and Taylor (2002), Levich (2017) and Cerutti et al. (2021) trace the first efforts to estimate the neutral band back to Keynes (1923).

2.3 Intermediary Asset Pricing Model

International finance literature has estimated $\underline{b}_{t,\tau}$ and $\overline{b}_{t,\tau}$ with either a counting approach, whereby values are such that 95% of observed financial deviations are contained, resembling the value-at-risk methodology; or linear (Branson, 1969), or non-linear in mean econometric techniques (Peel and Taylor, 2002; Juhl et al., 2006) that estimate thresholds beyond which $x_{t,\tau}$ reveals arbitrage opportunities. Both approaches may be seen as backward-looking as they determine the neutral band estimate based on all previously observed $x_{t,\tau}$. They also model the neutral band boundaries and arbitrage opportunities as being determined simultaneously. However, proceeding in this way overlooks the timing of decisions made by market participants.

Let us consider the following sequence of events, which more closely resembles the decision process within a financial institution aiming to take advantage of arbitrage opportunities. Before entering any arbitrage transaction, the market participant knows $\underline{b}_{t,\tau}$ and $\overline{b}_{t,\tau}$, possibly estimated by a risk-management department.[3] Therefore, risk management should estimate $b_{t,\tau}$ *before* any arbitrage transaction takes place at period t. This requirement, and the persistent deviations from CIP, make the last observed deviation from CIP the best starting point for producing the estimate $b_{t,\tau}$, as opposed to the *theoretical* zero value.

To formalize this decision process, consider Du, Hébert, and Huber (2022) (DHH)'s intermediary asset pricing model with a regulatory constraint based on He and Krishnamurthy (2018).[4] In this model, the (risk) manager of an intermediary maximises Epstein and Zin (1989) preferences over consumption and a portfolio of assets indexed by $i \in I$ facing (among others) a period-by-period regulatory constraint

$$\sum_{i \in I} k^i |\alpha_t^i| < 1, \tag{2.5}$$

where k^i is the asset i specific weight, α_t^i is the intermediary's holding of asset i in period t. In this model, long and short positions are allowed, hence the need for the $|\cdot|$ function. As explained by DHH, the portfolio allocation constraint (2.5) captures macroprudential limits, such as leverage ratios and risk-weighted capital. The equilibrium condition for α_t^i is given by

$$E_t\left[\exp\left(m_{t+1}\right)\left(R_{t+1}^i - R_t^b\right)\right] = \lambda_t^{RC} k^i sgn(\alpha_t^i) \tag{2.6}$$

where $E_t[\cdot]$ is the expectation conditional on the manager of the intermediary's wealth and information available at period t, m_{t+1} is the (log-)stochastic discount factor (SDF), R_{t+1}^i is the gross return on asset i, R_t^b is the gross return on the intermediary's debt between periods t and $t+1$, λ_t^{RC} is the scaled multiplier on the regulatory constraint and $sgn(\cdot)$ is the sign function.

Consider the arbitrage strategy on deviations from CIP (i.e. a strategy to profit from the cross-country basis when $x_{t,\tau} \neq 0$). If the intermediary can obtain funding with an interest rate of $y_{t,\tau}$, applying (2.6) will result in

$$E_t\left[\exp\left(m_{t+1}\right)\right] R_t^b \left|1 - \exp\left(-x_{t,\tau}\right)\right| = \lambda_t^{RC} k^c \tag{2.7}$$

where k^c is the weight on the cross-currency basis arbitrage. A few remarks on (2.7) are in order: as pointed out by DHH, the absolute value of $x_{t,\tau}$ is a key input to estimate the (shadow) cost of (2.5); $x_{t,\tau}$ is non-stochastic conditional on information available in period t since all transactions related to it take place in period t; and there is a symmetry implied by the absolute value.

Campbell (1993) and Du et al. (2022) show that by combining (2.6) and (2.7), and log-linearising the result, it is possible to derive an expression for revisions in future investment opportunities which, in turn, responds to revisions in the cross-currency basis

$$\left(E_{t+1} - E_t\right) \log R_{t+1+j}^w = \left(E_{t+1} - E_t\right)\left[\log R_{t+j}^b + (k^c)^{-1} x_{t+j,\tau} + V_{t+j}\right] \tag{2.8}$$

[3] In the aftermath of the GFC, this has been a requirement from financial authorities in many jurisdictions. On margin requirements, for example BIS (2019) in paragraph 20.12 states that: "Non-centrally cleared derivatives will often be exposed to a number of complex and interrelated risks. Internal or third-party quantitative models that assess these risks in a granular form can be useful for ensuring that the relevant initial margin amounts are calculated in an appropriately risk-sensitive manner. Moreover, current practice among a number of large and active central counterparties is to use internal quantitative models when determining initial margin amounts". See IMF (2020) for another example.

[4] See the internet appendix of DHH for further details.

where R^w_{t+1} is the gross return on the manager portfolio and V_t contains time-varying variance of the assets and the currency basis. Assuming homoskedasticity, DHH arrive to (2.8) with $V_t = 0$.

Expression (2.8) motivates the estimation of the neutral band below. In particular, one-period revisions determine, as mentioned, arbitrage opportunities from CIP deviations. Also, revisions on the variance of the basis play a role in revisions of both the basis and investment opportunities.

Revisions require a one-step-ahead prediction for $x_{t,\tau}$, and a regulatory constraint in place that guards against risk. Here lies the main difference between this paper and previous work. Instead of assuming that $x_{t,\tau}$ and $b_{t,\tau}$ are determined simultaneously, I obtain a measure of "large enough" profitable deviations through a (one-step ahead) predictive distribution with time-varying variance. This approach synthesises both the counting and econometric approaches previously described.

2.4 Data

The data on deviations from CIP that I use in the empirical analysis are the latest update of what is used in Du and Schreger (2016) and Du et al. (2018), and provided by the authors. This database, which includes different tenors for sovereign interest rates, offers several advantages. In particular, transformations are not required to obtain returns in annual terms and allows for easy comparison and discussion on previous work.

I focus on deviations from CIP with 3-month maturity assets, $\tau = 3$ months. A number of previous studies that estimate $b_{t,\tau}$ have used this tenor (Frenkel and Levich, 1975, 1977; Peel and Taylor, 2002; Du et al., 2018; Cendese et al., 2021; Cerutti et al., 2021). Cerutti et al. (2021) focus on the 3-month LIBOR basis, however. I use expression (2.4) to compute deviations from CIP where $y_{t,\tau}$ is the interest rate on the 3-month US Treasury Note in USD, $y^*_{t,\tau}$ is the interest rate on the 3-month Gilt in GBP and $\rho_{t,\tau}$ is the 3-month forward premium measured in GBP per USD.

Deviations from CIP are displayed in Figure 1. The daily sample covers from 3 January 2000, as in Du et al. (2018), to 21 March 2021 for a total of 5,527 observations. Summary statistics are contained in Table 6. The sample does not include a further major stress event for the GBP commonly referred to as the "mini-budget debacle" in September 2022. This period encompasses several global financial stress episodes, most notably the global financial crisis (2007-2009), the sovereign debt crisis in Europe (2010-2011), and the COVID-19 pandemic (2020). As for stress episodes specific to the GBP, it includes the Brexit referendum and the ensuing negotiations (2016-2018).

Figure 1 shows that, from 2000 to 2006, deviations from CIP were around a mean with lower volatility than in the GFC, which started in 2007 and ended in 2009. This markedly distinct behaviour of $x_{t,\tau}$ relates to financial stress events. The Figure also reveals how $x_{t,\tau}$ only gradually returned to pre-GFC levels during 2009, and exhibited high volatility from mid-2015 to the end of 2017.

Prices have persistently deviated from their behaviour prior to the GFC. The literature identifies three classes of factors as an explanation, as Cerutti et al. (2021) summarises: (i) risk appetite, (ii) monetary policies and (ii) financial regulations. In the following sections, I focus on discussing three specific periods around the GFC, as outlined in Cerutti et al. (2021); Du et al. (2022): pre-GFC 2000-2006, GFC 2007-2009, post GFC 2010-2021.

The sign of the majority of observations of $x_{t,\tau}$ is negative, which means that the US Dollar market interest rate is consistently lower than the synthetic interest rate. Figure 2 displays the histogram with a long left-tail. This is not particularly thick, however. The 95% of the density below the curve corresponds to deviations of -53.1 basis points or higher.

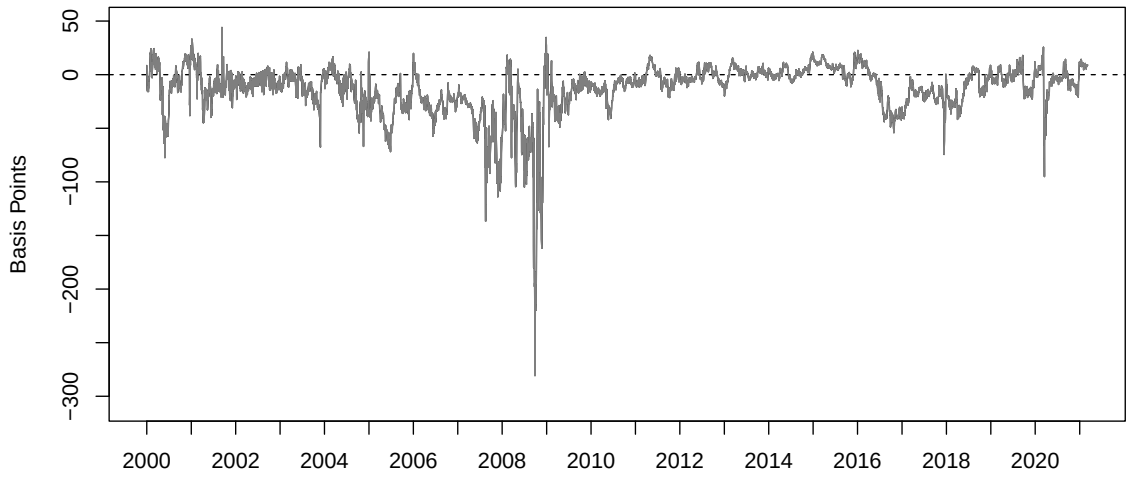

Figure 1: Deviations from CIP, $x_{t,\tau}$, in basis points computed as in (2.4) for the 3-month sovereign basis. Positive (negative) values of $x_{t,\tau}$ correspond to the cases where the USD synthetic rate is smaller (larger) than the market USD rate. Sample: 3rd January 2000 to 21st March 2021. Source: Du and Schreger (2016) and Du et al. (2018).

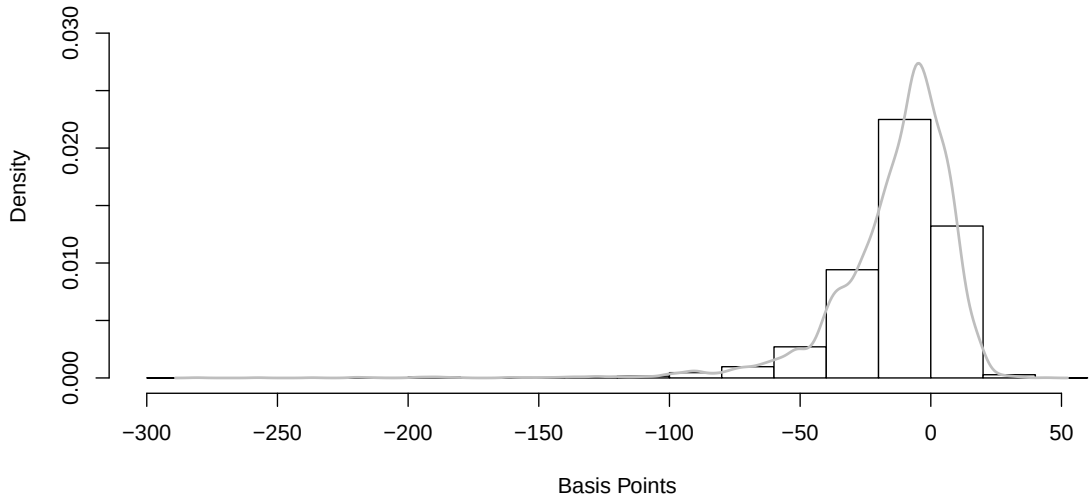

Figure 2: Density of deviations from CIP, $x_{t,\tau}$, in basis points computed as in (2.4) for the 3-month sovereign basis. Positive (negative) values of $x_{t,\tau}$ correspond to the cases where the USD synthetic rate is smaller (larger) than the market USD rate. Sample: 3rd January 2000 to 21st March 2021. Source: Du and Schreger (2016) and Du et al. (2018).

3 Econometric Methods

In this Section, I present the models used to estimate b_t as defined in Section 2 (i.e. one-step ahead predictive distributions). Since all assets with a return have the same 3-month horizon, I omit the

τ index from now on to ease notation. Formally, let $\mathcal{F}_t$ represent the information available up to time t. $\mathcal{F}_t$ contains observable data, as well as any binding regulatory constraint on each manager. Conditionally on the "counting" approach, and the timing of the risk managers decisions, I base the neutral band boundary estimates, $\bar{b}_t$ and $\underline{b}_t$, on the observed values of x_t that satisfy

$$p\left(x_t > \bar{b}_t|\mathcal{F}_{t-1}\right) = p\left(x_t < \underline{b}_t|\mathcal{F}_{t-1}\right) = 0.025, \tag{3.1}$$

where $p\left(\cdot|\mathcal{F}_{t-1}\right)$ is the predictive distribution for x_t conditional on available information at *t-1*, $\mathcal{F}_{t-1}$. Expression (3.1) is similar to that of Brunnermeier and Pedersen (2008) for margin calls. $\underline{b}_t$ and $\bar{b}_t$ are particular values of the estimated predictive distribution of x_t, $\widehat{p}\left(\cdot|\mathcal{F}_{t-1}\right)$. I define the neutral band as $b_t = \left[\underline{b}_t, \bar{b}_t\right]$.

The predictive distribution function associated with all the models I considered below is Gaussian

$$\widehat{p}\left(x_t|\mathcal{F}_{t-1}\right) = \mathcal{N}\left(\widehat{x}_{t|t-1}, \widehat{V}_{t|t-1}\right), \tag{3.2}$$

where the conditional mean forecast and the conditional variance are obtained from (3.5) and (3.6), and are given by

$$\widehat{x}_{t|t-1} = E\left(x_t|\mathcal{F}_{t-1}\right), \tag{3.3}$$

$$\widehat{V}_{t|t-1} = Var\left(x_t|\mathcal{F}_{t-1}\right). \tag{3.4}$$

To obtain these estimates, based on inspection of Figure 1 and on Du et al. (2022), I assume that $\{x_t\}_{t=1}^n$ is autoregressive of order 1, and add the (log) time-varying conditional variance h

$$x_t = \beta_0 + \beta_1 x_{t-1} + \exp(h_t/2)\varepsilon_t, \tag{3.5}$$

$$h_t = \theta_0 + \theta_1(h_{t-1} - \theta_0) + \theta_2 x_{t-1}^2 + \sigma_\eta \eta_t, \tag{3.6}$$

$$\begin{pmatrix}\varepsilon_t \\ \eta_t\end{pmatrix} \sim N\left[\begin{pmatrix}0 \\ 0\end{pmatrix}, \begin{pmatrix}1 & \rho \\ \rho & 1\end{pmatrix}\right]. \tag{3.7}$$

I group the parameters of the conditional mean and the conditional variance in $\beta = (\beta_0, \beta_1)'$ and $\theta = (\theta_0, \theta_1, \theta_2, \sigma_\eta, \rho)'$, respectively. The conditional variance follows a mean-reverting first-order autoregressive process with mean θ_0. The mean reversion of the variance is a well-established feature of financial variables. Another feature included in (3.7) is leverage, which corresponds to cases in which $\rho \neq 0$.

The rest of the section details the estimation of (3.5) and (3.6). I present four models, one that assesses non-linearities in mean (threshold autoregressive model), and three with autoregressive conditional variance (generalised autoregressive conditional heteroskedasticity model, stochastic volatility model and stochastic volatility with leverage model).

Threshold Autoregressive (TAR) Model

The method previously used in the literature for estimating b_t, by Peel and Taylor (2002) and Juhl et al. (2006), is the threshold autoregressive (TAR) model from Tong (1990) and Granger and Teräsvirta (1993). Therefore, this is the natural candidate for a benchmark model. I can re-state the TAR as a particular case of (3.5) and (3.6) with $\beta^T = \left(\beta_0^D, \beta_1^D, \beta_0^M, \beta_1^M, \beta_0^U, \beta_1^U\right)'$ and $\theta^T = \theta_0^T$,

$$x_t = \begin{cases} \beta_0^D + \beta_1^D x_{t-1} + \left(\theta_0^T\right)^{1/2}\varepsilon_t & \text{if } x_{t-1} \leq \kappa^D, \\ \beta_0^M + \beta_1^M x_{t-1} + \left(\theta_0^T\right)^{1/2}\varepsilon_t & \text{if } \kappa^D < x_{t-1} < \kappa^U, \\ \beta_0^U + \beta_1^U x_{t-1} + \left(\theta_0^T\right)^{1/2}\varepsilon_t & \text{if } x_{t-1} \geq \kappa^U, \end{cases} \tag{3.8}$$

where κ^U and κ^D are the upper- and lower-threshold levels in (3.8). In the particular case where $\beta^T = (0, 1)'$, the TAR is consistent with a no-arbitrage condition (i.e. consistent with a Martingale process). It lacks, however, a time-varying variance component, a widely known feature of financial data and present in (2.8). I use the following models to address this.

Generalised Autoregressive Conditional Heteroskedasticity (GARCH) Model

Bollerslev (1986) introduced the generalised autoregressive conditional heteroskedasticity (GARCH) model, which is widely used in financial analysis, with important applications in value-at-risk (Ardia, 2008). It is relatively straightforward to estimate and interpret, and its properties have been thoroughly studied.

However, a caveat of the GARCH is that it cannot accommodate "shocks" to the conditional variance process. The GARCH may be consistent with no-arbitrage and is also a particular case of (3.5) and (3.6) with parameters to estimate $\beta^G = \left(\beta_0^G, \beta_1^G\right)'$ and $\theta^G = \left(\theta_0^G, \theta_1^G, \theta_2^G\right)'$,

$$x_t = \beta_0^G + \beta_1^G x_{t-1} + \left(h_t^G\right)^{1/2} \varepsilon_t, \tag{3.9}$$

$$h_t^G = \theta_0^G + \theta_1^G h_{t-1}^G + \theta_2^G x_{t-1}^2, \tag{3.10}$$

where (3.9) is the observation equation and (3.10) is the state equation. The parameter vector θ^G is restricted so that the estimate for the conditional variance, h_t^G, is positive.

Stochastic Volatility (SV) Model

The stochastic volatility (SV) model, in addition to its parsimony and intuitive interpretation, can represent a large class of Martingale processes (Shephard, 2015). The SV shares the same appealing features of the GARCH, but can accommodate shocks to volatility. The model, described in detail in Kim, Shephard, and Chib (1998), is given by

$$x_t = \beta_0^S + \beta_1^S x_{t-1} + \exp(h_t^S/2)\varepsilon_t, \tag{3.11}$$

$$h_t^S = \theta_0^S + \theta_1^S(h_{t-1}^S - \theta_0^S) + \sigma_\eta^S \eta_t, \tag{3.12}$$

where (3.11) is the observation equation, and (3.12) is the state equation with the restriction that $-1 < \theta_1^S < 1$. Both are particular cases of (3.5) and (3.6), with $\theta_2 = \rho = 0$. The objects to be estimated in the SV model are the parameter vectors $\beta^S = \left(\beta_0^S, \beta_1^S\right)'$ and $\theta^S = \left(\theta_0^S, \theta_1^S, \sigma_\eta^S\right)'$, and the latent process for the (log) time-varying conditional variance h_t^S.

Stochastic Volatility with Leverage (SVL) Model

By adding the parameter $\rho \neq 0$ in (3.7) to the model, the model can help in determining if there is a leverage effect in the cross-country basis. Here, leverage is defined as the effect of unexpected changes in the currency basis on the dynamics of the conditional variance. This generalization allows me to explicitly model the asymmetry of x_t, a feature displayed in Figure 2. A stochastic volatility with leverage (SVL) model is given by (3.11) and (3.12) with parameter vectors $\beta^L = \left(\beta_0^L, \beta_1^L\right)'$ and $\theta^L = \left(\theta_0^L, \theta_1^L, \sigma_\eta^L, \rho^L\right)'$, and conditional volatility process h_t^L.

4 Estimates and Diagnostics

I present two sets of estimates for each model. First, I carry out a full sample estimation, which allows us to analyse the specification for each model and assess in-sample fit. Second, I estimate forecasts using an expanding window sample to account for the real-time reassessing feature of $\underline{b}_t$ and $\overline{b}_t$. In this set, the training sample contains 360 observations, from 3 January 2000 to 21 May 2001. No events causing great or long-lasting financial stress were registered for the USD-GBP market during this period.

To conduct inference by means of the density forecast defined in (3.2), estimates for β, θ and h_t are needed. I estimate models through classical-statistics (TAR and GARCH) or bayesian-statistics (SV, SVL) methods and evaluate them with the corresponding in- and out-of sample statistics. Details on the posterior simulators, detailed in Appendix B. Before discussing the neutral band estimates, I provide a discussion of the parameter estimates for the full sample.

4.1 Full Sample Parameter Estimates

I estimate the TAR model by non-linear least-squares with an autoregressive component of order 1 and 3 regimes, as in Peel and Taylor (2002). As previously mentioned, this is the model used in the literature to obtain an econometric estimate of b_t. Therefore, it serves as a natural benchmark to assess the results from the set of models described above.

Table 1 presents the results from the non-linear least-squares in the TAR columns. The results suggest that both the low and mid regimes resemble a no-arbitrage condition, since $\hat{\beta}_1^L, \hat{\beta}_1^M$ are relatively close to 1. However, Augmented Dickey-Fuller (unit root) tests for x_t in all regimes reject the Null Hypothesis of a unit root. The estimated thresholds, $b_t = \left[\kappa^D, \kappa^U\right]$, reflect that most observations of x_t are negative. If $x_t < -24.15$ basis points, the estimated parameters correspond to column "Lower" in the table. Similarly, if $x_t > 4.73$ basis points, the model with parameters shown in column "Upper" fits the data. The estimates for β_1 are in line with the non-explosive behaviour of x_t displayed in Figure 1 since they are all smaller than 1.

Table 1 presents the full sample estimates for the GARCH model. All estimates of the variance process are positive, but $\hat{\theta}_0 = 0.231$ with a standard error of 0.193 is not centred away from zero. While it is possible to refine these results, the mean estimates suggest that the conditional variance plays a heavily stabilising role in the dynamics of x_t. To see this, as discussed in Ardia (2008), the GARCH can be rewritten as an autoregressive-moving average of order (1,1) process with autoregressive parameter $\theta_1^G + \theta_2^G$ and moving average parameter $-\theta_2^G$. This means that the GARCH model needs a large value associated to the autoregressive term in the conditional variance to fit the data $\hat{\theta}_1^G + \hat{\theta}_2^G = 0.998$. The estimate for autoregressive parameter $\hat{\beta}_1 = 0.974$ is close to 1, suggesting a high level of persistence of x_t. It is easy to reconcile this feature with the observed values of x_t displayed in Figure 1, since they display mean-reversion.

		TAR regimes					
		Lower $\kappa^D = -24.15$	Mid	Upper $\kappa^D = 4.73$	GARCH	SV	SVL
β							
	β_0	-2.606	-0.551	1.641	-3.822	-0.090	-0.105
		(0.371)	(0.171)	(0.535)	(1.778)	(0.040)	(0.042)
	β_1	0.929	0.931	0.754	0.974	0.979	0.978
		(0.007)	(0.015)	(0.047)	(0.004)	0.003	(0.003)
θ							
	θ_0	6.811	6.811	6.811	0.231	2.230	2.370
					(0.145)	(0.193)	(0.206)
	θ_1				0.125	0.975	0.987
					(0.032)	(0.005)	(0.003)
	θ_2				0.873	-	-
					(0.035)	-	-
	σ_η					0.096	0.041
						(0.012)	0.006
	ρ						-0.063
							(0.049)

Table 1: Estimates for (3.5) and (3.6). Standard errors in parenthesis. Non-linear least squares estimates for the TAR model parameters β^T and lower and upper threshold estimates κ^L, κ^U. θ_0 is the error unconditional variance. Maximum likelihood estimates for the GARCH model parameters, β^G, θ^G. SV and SVL model parameters: posterior mean of β^S, θ^S, β^L, θ^L; Monte Carlo standard error in parenthesis. Details are contained in Appendix B.

I estimate the SV through Markov Chain Monte Carlo (MCMC) methods proposed by Kim et al. (1998). An advantage of modelling volatility as a stochastic time-varying process is the model's ability to identify periods of sudden increases in uncertainty. A caveat of this method, when producing forecasts, is that, every time a new observation is collected, the MCMC must be implemented to obtain the posterior distribution of h_t^S. Table 1 shows that all elements of β^S and θ^S are centered away from zero. The autoregressive estimate $\hat{\beta}_1 = 0.979$ is close too 1, similar to that of the GARCH model. The autoregressive term for the variance process $\hat{\theta}_1^S = 0.975$ is also close to 1, confirming the highly persistence of the variance process. This is in line with mean-reverting behaviour of x_t and the well established fact that variance in assets is stationary.

The estimates also show that the standard deviation $\hat{\sigma}_\eta^S = 0.096$ is relevant for modelling the time-varying conditional volatility. Should this estimate turn out to be (close to) zero, the GARCH would be a more parsimonious way to fit the behaviour of x_t. However, this is not the case. This estimate underlies how unexpected changes in the variance process play role in explaining the dynamics of the variance of x_t and its conditional mean level.

The Bayesian estimates for the SVL are obtained through MCMC as in Omori et al. (2007). The last column in Table 1 displays a summary of the estimates for β^L and θ^L for the full sample. Estimates of β are very similar to those of the SV model, but those of θ differ. $\hat{\sigma}_\eta^L = 0.041$ is less than half $\hat{\sigma}_\eta^S$. The estimate for the leverage parameter $\hat{\rho}$ is negative, which would suggest the presence of the leverage effect, but center away from zero by 1.3 standard deviations.

4.2 Neutral Band Estimates

Estimates for b_t, obtained through one-step ahead predictive distributions in a expanding window, reflect the general features of the constant variance model (TAR) and the time-varying variance models (GARCH, SV, and SVL). Estimates from the TAR model are wide and seem to adjust slowly to new information. As I discuss below, this slow adjustment is difficult to reconcile with both theory and financial stress events observed in the sample. In contrast, estimates from the GARCH, SV and SVL can reflect changes in uncertainty, which may cause changes in transaction costs (e.g. margins) and risks associated with CIP.

The GARCH, SV, and SVL estimates imply arbitrage opportunities that are dispersed through time, a feature more closely associated with the mechanics of FX markets. The clustering of the suggested arbitrage opportunities by the TAR is implausible when arbitrage is present and when FX markets are liquid. Such clustering is also undesirable from a forecasting performance perspective as it reveals dependence in coverage (Christoffersen, 1998). A good forecasting model would yield independent forecast errors (i.e., observations of x_t outside b_t should be dispersed in time).

The slow adjustment and the width of the b_t estimates from the TAR seem to be a result from its constant variance, notwithstanding the expanding window estimation approach. Results reveal that the TAR's b_t estimates for the GBP-USD cross are notably wide, as displayed in Figure 3, along with x_t. As previously mentioned, observed values of x_t at period t outside b_t (marked with a vertical line) are clustered. These estimates fail to suggest any arbitrage opportunity between early 2009 and late 2017. However, this contradicts the documented increases in uncertainty and transaction costs in several episodes in this subset of the sample. The GBP-USD cross experienced episodes of considerable financial stress during the sovereign debt crisis in Europe, between 2010 and 2012, and the "Brexit" referendum in mid-2016. The TAR model fails to recognise the latter as a period where the width of b_t changes.

Observed values of x_t outside b_t obtained from the GARCH, SV and SVL, displayed in Figures 4-6, appear short-lived and are more dispersed through time. Variation in the magnitude of b_t estimates also reflects the referred financial stress events. Comparing the results from these models, GARCH estimates of b_t prior to mid 2007 suggest fewer arbitrage opportunities than the SV and SVL. While financial stress in this period was milder when compared with the GFC, there are some events that might have caused increases in uncertainty. Three examples are the war in Afghanistan and Iraq in 2003; the Federal Reserve tightening cycle in 2004-2006; and bond market dislocations in Europe due to fiscal deficits in Germany and France in 2002-2005.

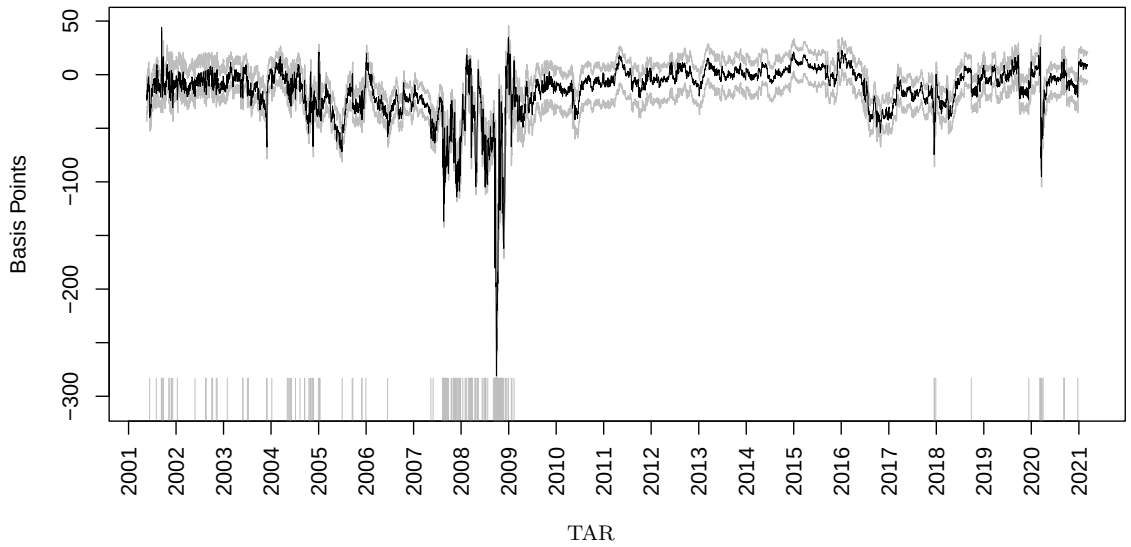

Figure 3: Black: Deviations from CIP, x_t, in basis points. Grey: TAR model Neutral band estimates for GBP-USD, defined as the 97.5% and 2.5% quantile estimate of the one step-ahead predictive distribution $\hat{p}(x_t|\mathcal{F}_{t-1})$. The grey vertical markers are occurrences of $x_t > \hat{\bar{b}}_t$ or $x_t < \hat{\underline{b}}_t$ in basis points. Source: Own estimates and Du and Schreger (2016) and Du et al. (2018).

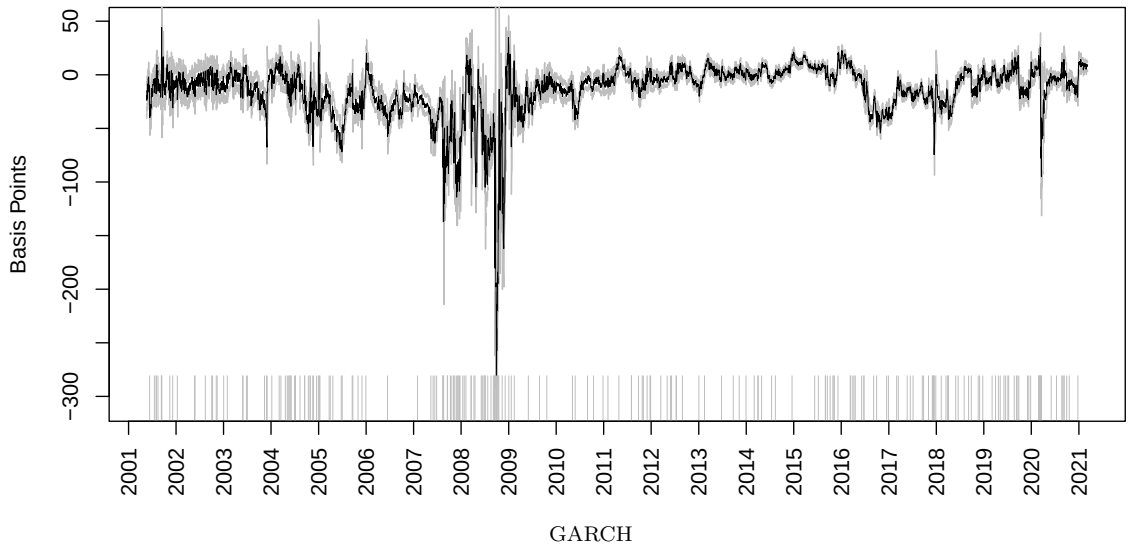

Figure 4: Black: Deviations from CIP, x_t, in basis points. Grey: GARCH model Neutral band estimates for GBP-USD, defined as the 97.5% and 2.5% quantile estimate of the one step-ahead predictive distribution $\hat{p}(x_t|\mathcal{F}_{t-1})$. The grey vertical markers are occurrences of $x_t > \hat{\bar{b}}_t$ or $x_t < \hat{\underline{b}}_t$ in basis points. Source: Own estimates and Du and Schreger (2016) and Du et al. (2018).

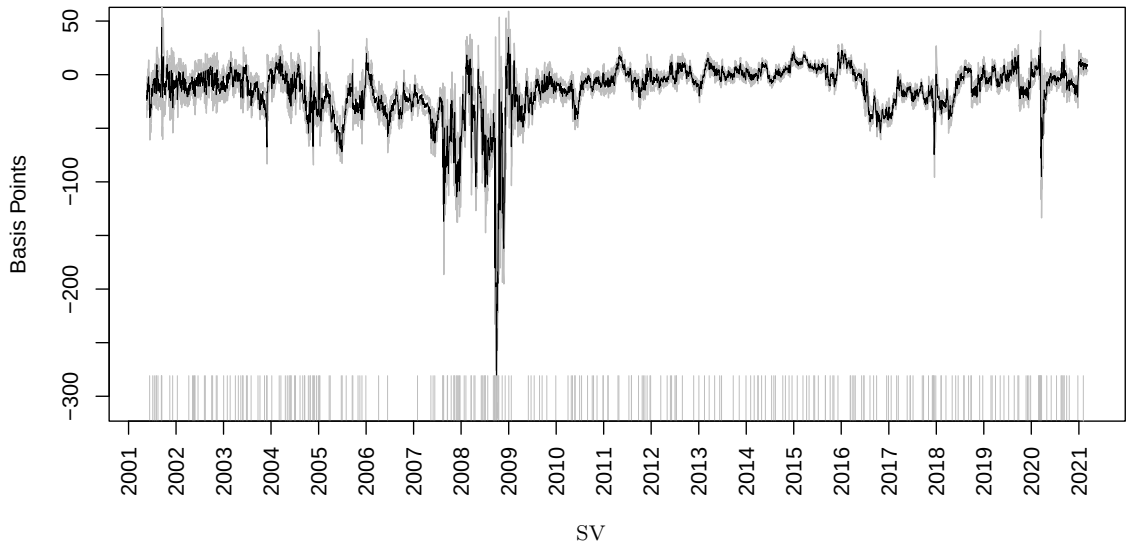

Figure 5: Black: Deviations from CIP, x_t, in basis points. Grey: SV model Neutral band estimates for GBP-USD, defined as the 97.5% and 2.5% quantile estimate of the one step-ahead predictive distribution $\hat{p}(x_t|\mathcal{F}_{t-1})$. The grey vertical markers are occurrences of $x_t > \hat{\bar{b}}_t$ or $x_t < \hat{\underline{b}}_t$ in basis points. Source: Own estimates and Du and Schreger (2016) and Du et al. (2018).

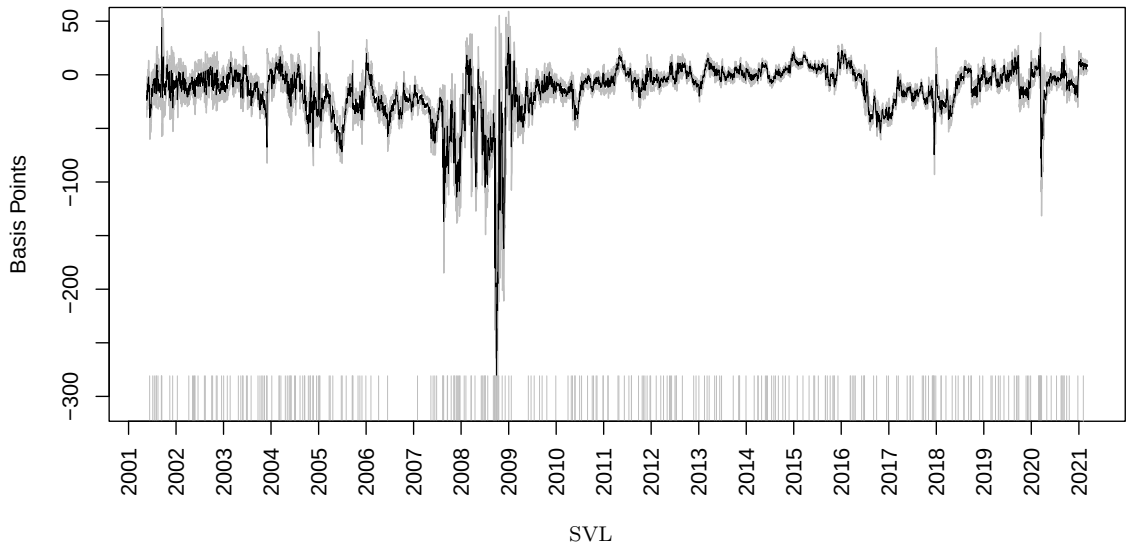

Figure 6: Black: Deviations from CIP, x_t, in basis points. Grey: SVL model Neutral band estimates for GBP-USD, defined as the 97.5% and 2.5% quantile estimate of the one step-ahead predictive distribution $\hat{p}(x_t|\mathcal{F}_{t-1})$. The grey vertical markers are occurrences of $x_t > \hat{\bar{b}}_t$ or $x_t < \hat{\underline{b}}_t$ in basis points. Source: Own estimates and Du and Schreger (2016) and Du et al. (2018).

Table 2 presents a summary of the estimates for the width of b_t. For ease of exposition, I divide the following discussion into three periods, as in Cerutti et al. (2021); Du et al. (2022): pre-GFC (2000-2006), during-GFC (2007-2009) and post-GFC (2010-2021).

To put these estimates in context, previous measures of the neutral band for the GBP-USD cross found in the literature somewhat resemble those for the pre-GFC period. Clinton (1988) and Frenkel and Levich (1975) estimate transaction costs of ±15 basis points, translating to $\bar{b}_t - \underline{b}_t = 30$ basis points. Frenkel and Levich (1977) in turn estimate transactions costs in the range from 12.6 to 100.3 basis points, conditionally on the stress observed in FX markets. This translates to $25.2 \leq \bar{b}_t - \underline{b}_t \leq 200.6$ basis points. The estimated b_t in these papers, and this work, are well below early estimates from Keynes and Enzig, but align with the development of the relevant markets and communication technology.[5]

The results for the pre-GFC period indicate that the highest estimates for the width of b_t between 31 and 167 basis points. Estimates are similar within the time-varying variance class of models, and are between 8 (SVL) and 167 basis points (GARCH), whereas TAR estimates are between 25 and 35 basis points. During the GFC, the estimated lowest value of the width of b_t ranges between 5 (SVL) and 29 basis points (TAR). However, the highest value of the range varies more for each model. The TAR estimates 37 basis points, whereas the GARCH, SV, and SVL estimate 290 and 191 and 233 basis points, respectively. Therefore, the GARCH and SVL have a broader range in a period of particularly acute financial stress.

Estimates from the TAR for the post-GFC period show that maximum values for the neutral band closely resemble those obtained for the period during the GFC. This suggests that determinants of b_t (e.g. uncertainty and transaction costs) are similar to those prevailing at the peak of the GFC. The GARCH and SV models, in turn, estimate a range of 3-102 basis points, higher than pre-GFC, but nowhere near the peaks attained then. As discussed in Christoffersen (1998), time-varying predictive distributions are an appealing feature of time-varying volatility models. However, this feature is also present in the SVL, albeit to a lesser extent. The range is small, compared to those from the GARCH and SV.

Model	Estimated range for $\bar{b}_t - \underline{b}_t$			Marginal Likelihood
	Pre-GFC	During-GF	Post-GFC	
TAR	25 to 31	29 to 37	27 to 34	-18430
GARCH	17 to 167	12 to 290	6 to 101	-15809
SV	9 to 107	7 to 191	3 to 102	-14762
SVL	8 to 116	5 to 233	4 to 95	-14951

Table 2: Estimated range for the neutral band width, $\bar{b}_t - \underline{b}_t$ in basis points, for the GBP-USD FX market in each period. Marginal likelihood for each model, the higher marginal likelihood (4.1) corresponds to the model with best fit. Pre-GFC: 2000-2006, During-GFC: 2007-2009 and Post-GFC: 2010-2021.

4.3 Model Diagnostics

The evidence presented in the previous section suggests that, in general, the TAR estimates a neutral band that changes slowly. The upside, in terms of risk management, is that these estimates may provide a more conservative b_t. However, this comes with an opportunity cost in terms of missed arbitrage opportunities. The TAR model fails to account for financial stress periods, which should reflect in increases in the width of b_t. In contrast, the GARCH, SV, and SVL (time-varying variance models) can account for changes in financial conditions, and hence b_t is consistent with short-lived arbitrage opportunities.

This analysis suggests that time-varying variance models are better candidates for producing estimates of b_t consistent with economic intuition. But the differences across estimates of the neutral band from GARCH, SV and SVL in Figures 4-6 models are rather subtle. To determine which model is more suitable to estimate b_t, I conduct a formal model evaluation in the rest of this section. I

[5] Keynes (1923) asserted that deviations from CIP would not be corrected entirely and instantaneously by transactions in the FX futures market since prices in London and New York took days to synchronize and supply of funds for market participants was relatively scarce.

use the marginal likelihood statistic to assess in-sample fit, the coverage ratio tests and the log-score comparisons for density forecast evaluation.

Marginal Likelihood

The posterior densities for the parameters and the latent processes allows to compute the marginal likelihood, ML_n and assess the fit of each model for the full sample of size n. Several methods have been proposed to compute the ML_n statistic for models estimated by simulation methods. Here, I follow Kim et al. (1998)

$$ML_n = \sum_{t=1}^{n} \frac{1}{M} \sum_{j=1}^{M} \ln \widehat{p}\left(x_t | x_{t|n}^{(j)}, V_{t|n}^{(j)}\right), \tag{4.1}$$

where $\mathcal{F}_n$ is the full information set, and M is the number of elements in the posterior distribution of $\hat{x}_{t|n}$. The last column in Table 2 summarises the estimates for each model. Direct comparison suggests that the TAR model does not fit the data as well as the GARCH model. The marginal likelihood al suggests the in-sample fit from the SV is better that that of the SVL, which in turn is better than that of the GARCH and the TAR. The difference in ML_n between the models with stochastic volatility and the TAR and GARCH models is non-negligible. The latent process estimates for h_t^G, h_t^S and h_t^L seem to play an important role for each model in fitting the data, particularly those with a stochastic component.

Coverage Ratio Tests

To assess the accuracy of the forecasting intervals, I use the coverage ratio likelihood ratio tests from Christoffersen (1998). The list of desirable features of forecast errors includes congruence with the confidence levels or credibility intervals (tested using LR_{uc}), that they are independent (tested using LR_{ind}), and that they have the correct coverage conditional on being independent (tested using LR_{cc}). These tests focus on whether the band contains 95% of the realized values for x_t, while penalising clustering and dependence among the realized values outside the band. These are desirable features from the forecasting perspective and in assessing arbitrage opportunities.

Table 3 display the results, along with the empirical coverage rate. The null hypothesis for LR_{uc} is correct coverage. The GARCH and SV provide the correct coverage, in contrast to the TAR and the SVL. The latter two models estimate features that the data do not support (non-linearities in mean and leverage). This aligns with the poor predictive properties of TAR models documented, for example, in Dacco and Satchell (1999). In turn, the SVL may be over-fitting the data, as it estimates an additional feature than the GARCH and SV, and may induce a slow adjustment of the forecasts to new information.

The null hypothesis for the LR_{ind} test statistic is that forecast errors outside b_t are independent,and the alternative is that they follow a Markov process of order 1. In this case, the clustering of forecast errors displayed by the TAR allows us to reject the null hypothesis, while the GARCH and SV display independent errors. The test also rejects the null hypothesis for the SVL. This suggests that autocorrelation exists in the forecast errors from this model. Finally, results from the LR_{cc} suggest that they hold jointly only for the SV. The slow adjusting neutral band width induced by the non-linear mean of the TAR, the slow changing conditional variance of the GARCH, and the SVL over-fitting seem to translate into incorrect coverage.

Model	Empirical Coverage	LR_{uc}	LR_{ind}	LR_{cc}
TAR	96.32	20.766	240.467	261.309
GARCH	95.56	3.596	9.934	13.620
SV	94.69	1.005	3.742	4.857
SVL	94.32	4.754	5.159	10.030
Critical values (95%)		3.841	3.841	5.991

Table 3: Coverage ratio tests and empirical coverage. The empirical coverage of each model is defined as the number of observations that fall within the neutral band. Christoffersen (1998) tests for the congruence of the theoretical 95% and the percentage of forecast errors, LR_{uc}, the test for independence of the forecast errors, LR_{ind}, and the test for the correct coverage conditional om the forecasts error being independent LR_{cc}. The null hypothesis is correct coverage. Critical values for the 95% are obtained from the χ^2_r distribution with r degrees of freedom.

Log-Score

To strike a balance between having a conservative b_t, and ensuring precision in terms of identifying financial stress periods, I conduct a density forecast evaluation through a pair-wise log-score statistic averaged over time, as suggested in Elliott and Timmermann (2016). To do this, first I estimate the log-score, LS_t, for each forecast to obtain the probability that the realised value of x_{t+1} is contained in the forecast distribution. Then I compute a log-score statistic, Q_{LS}. In particular

$$\begin{aligned} LS_t &= \frac{1}{M}\sum_{j=1}^{M} \ln \widehat{p}\left(x_t | x^{(j)}_{t|t-1}, V^{(j)}_{t|t-1}, \mathcal{F}_{t-1}\right), \\ Q_{LS} &= \frac{1}{n_f - 1}\sum_{j=1}^{n_f} LS_j, \end{aligned} \tag{4.2}$$

where M is the number of draws from the predictive distribution, $\widehat{p}(\cdot)$, and n_f is the number one-step ahead forecasts made. Even though I obtain the parameters of the TAR and GARCH models through classical estimation methods, it is possible to compare density forecasts as long as a time series for LS_t is recovered (Geweke and Amisano, 2010). Finally, I use Diebold-Mariano tests to determine if the log-score statistic differs across models. That is, the test statistic is the difference between pairs Q_{LS}, contrasted with the null hypothesis of equality, distributed as a student-t.

Results are presented in Table 4. A salient feature of the left-hand side block in the table is the low value of Q_{LS} for the TAR. Indeed, the log-score statistic for the time-varying conditional variance models is at -3.377 for the TAR and -2.831 for the GARCH. The SV has the highest Q_{LS} statistic among the analysed models. The formal pair-wise comparison between models, contained in the right block of the table, confirms that all alternative models to the TAR display better forecasting performance. In all cases, the positive t-statistic rejects the null hypothesis of equal Q_{LS}.

The pair-wise comparison between the time-varying variance models suggests that the SV offers superior forecasting performance. Several features inherent to the SV may explain this result. When comparing the SV with the GARCH, the inclusion of a stochastic term in the conditional variance dynamics gives the SV an additional degree of flexibility to replicate sudden, and temporary, large changes in the financial uncertainty around x_t. Additionally, this stochastic term provides more flexibility than the lagged term x_t included in the GARCH, which induces a slower pace in adjustments for the conditional variance.

The comparison between the SVL and the GARCH confirms, to some extent, the over-fitting of the former. While the in-sample evaluation suggests that the SVL had the best fit, both fail to produce the correct coverage. This may be due to the previous values of x_t playing an important role in

the GARCH (through the autoregressive moving average-like representation discussed above) and the SVL (through over-fitting). In other words, forecasts produced by these models adjust slowly to new information, but this is not a feature of the time series for x_t.

The comparison between models with a stochastic volatility process yields the SV as a superior candidate to estimate the neutral band. This may be due to its higher degree of flexibility to accommodate sudden, but temporary, increases in volatility, whereas estimates of the neutral band for the SVL may be slowly adjusting. This is an important consideration that risk management and forecasting precision share: it is undesirable to have forecasts that are autocorrelated with previous forecasts. In sum, the SV's ability to accommodate short-lived sudden changes complements the desirable feature of including a time-varying variance from the GARCH and is more flexible than that of the SVL.

Model	Q_{LS}	Accuracy	*t-stat*	p-value
		GARCH vs TAR	0.545	2.200×10^{-16}
TAR	-3.377	SV vs TAR	0.603	2.200×10^{-16}
GARCH	-2.831	SVL vs TAR	0.598	2.200×10^{-16}
SV	-2.774	SV vs GARCH	0.057	3.429×10^{-8}
SVL	-2.779	SVL vs GARCH	0.052	7.298×10^{-8}
		SV vs SVL	0.005	0.002

Table 4: Log-score comparison. The Q_{LS} statistic (4.2) summarises the one-step-ahead forecast precision of each model. Large precision corresponds to large Q_{LS}. The *t-stat* is the Diebold-Mariano test statistic that compares if two models have the same Q_{LS} statistically. The null hypothesis is: two models have the same Q_{LS}. Heteroskedasticity-Autocorrelation robust standard errors are used in the comparison.

5 Neutral Band Estimate for an Emerging Market Economy

The covered GBP-USD market is isolated from risks that are only associated with emerging market economies (EMEs). However, the approach I propose to estimate the neutral band is not conditional on whether there are only "reserve currencies" involved in the trade. In this section, I present the estimates of the neutral band for the covered Mexican Peso (MXN) US Dollar market.

Mexico is an EME with a floating exchange rate regime, an independent central bank and an inflation target framework for monetary policy. However, the Mexican economy and FX market face additional risks to those of the GBP-USD market. Global liquidity and additional risk premia, for example, have been found to determine deviations from CIP for the MXN-USD Hernandez (2014); Bush and López Noria (2021). The MXN-USD market is among the largest for an EME, only behind the Chinese Yuan and the Indian Rupee according to the BIS (2022).

I use the data on sovereign 3-month basis for the covered MXN-USD provided by Du and Schreger (2016) and Du et al. (2018), with daily observations from 2nd September 2002 to 9th March 2021. This results in a total of 4,817 observations displayed in Figure 7, with summary statistics in Table 6 in the Appendix. I estimate the same set of models (TAR, GARCH, SV, SVL) and follow the same approach described in the previous sections.

The results of the estimation, contained in Table 5, show that all models with a time-varying variance have a better fit with the data, and the SV model is the best among those. The models with stochastic volatility have better forecast performance than the GARCH model, but evidence is not conclusive on which has higher precision. The only distinctive feature between the SV and the SVL models is the conditional coverage likelihood ratio test, which suggests that the SV has better coverage. Further details of the estimation are contained in Tables 7-9 in Appendix A.

Table 5 contains the estimated range for the neutral band. The features that make the time-varying variance models more suitable are clearly displayed. The estimates obtained with TAR model are similar across time, while those for the GARCH, SV and SVL change with financial stress. The bands estimated from the GARCH model are markedly and statistically different from the those from the SV and SVL. Figures 8-11 in Appendix A display the neutral band estimates. However, the similarities

between the SV and SVL estimates are not statistically or economically significant. Details on the Diebold-Mariano tests can be found in Table 9 in Appendix A.

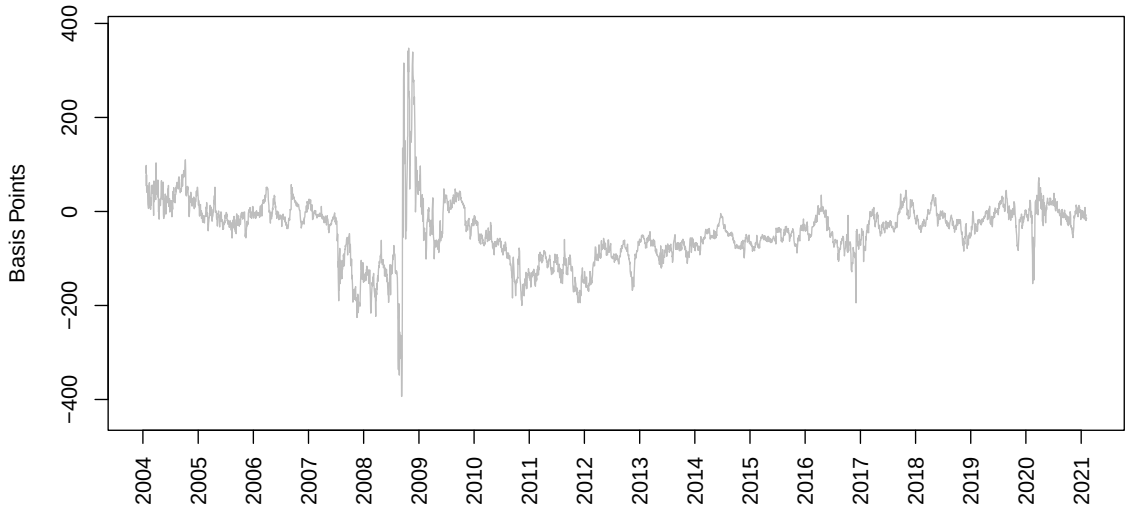

Figure 7: Deviations from CIP, $x_{t,\tau}$, in basis points computed as in (2.4) for the 3-month sovereign basis. Positive (negative) values of $x_{t,\tau}$ correspond to the cases where the USD synthetic rate is smaller (larger) than the market USD rate. Sample: 2nd September 2002 to 9th March 2021. Source: Du and Schreger (2016) and Du et al. (2018).

Model	Estimated range for $\overline{b}_t - \underline{b}_t$			Marginal Likelihood	Q_{LS}	Empirical Coverage
	Pre-GFC	During-GF	Post-GFC			
TAR	53 to 75	54 to 79	54 to 71	-19372	-4.131	97.08
GARCH	21 to 128	23 to 501	15 to 220	-17404	-3.588	95.55
SV	9 to 133	13 to 536	9 to 214	-16599	-3.536	94.82
SVL	9 to 134	13 to 543	9 to 205	-16675	-3.537	94.57

Table 5: Estimated range for the neutral band width for the covered MXN-USD, $\overline{b}_t - \underline{b}_t$ in basis points, for the GBP-USD FX market in each period. Marginal likelihood for each model, the higher marginal likelihood (4.1) corresponds to the model with best fit. Pre-GFC: 2000-2006, During-GFC: 2007-2009 and Post-GFC: 2010-2021. The Q_{LS} statistic (4.2) summarises the one-step-ahead forecast precision of each model. Large precision corresponds to large Q_{LS}. The empirical coverage of each model is defined as the number of observations that fall within the neutral band.

6 Concluding Remarks

In this paper, I argue that estimates of the neutral band around deviations from CIP, obtained from the forecast of an econometric model with time-varying variance, are intuitive and consistent with episodes of financial stress. The structural change observed in CIP after the great financial crisis in the covered GBP-USD FX market exposed some limitations of the previous estimates of the neutral

band. Changes in regulation requires a continuous monitoring of liquidity by risk-manager, deviations from CIP are no longer short-lived and the width of the neutral band should vary when risk changes. These are features that previous estimates do not account for, but they can be addressed by using a one step ahead forecast from a stochastic volatility model. I applied the approach to the MXN-USD currency cross and found that results are maintained, despite the different factors behind changes in this market.

The results from model evaluation suggest the stochastic volatility model, which is consistent with no-arbitrage, yields superior neutral band estimates. Several features suggest that these estimates have better properties than the alternatives. The model is able to replicate the past behaviour of deviations from CIP and displays a superior forecast performance. Like all models with time-varying variance, it estimates a wider neutral band in periods of financial stress. The neutral band is wider during the great financial crisis than in the rest of the sample, whereas previously used models estimate similar values for the January 2001 to March 2021 period. Finally, the neutral band estimate implies arbitrage opportunities scattered through time, hence, short-lived. In contrast, rival models imply either clustered or non-existent arbitrage opportunities over long periods of time. Both features are hard to reconcile with previous findings and covered GBP-USD FX market dynamics. The results for the MXN-USD mirror those for the GBP-USD.

There are some caveats in interpreting the results that require some consideration. Although it is possible to assess the in-sample fit of the models, estimates of the neutral band are based on the one-step ahead predictive distribution of a latent process. As such, it is not possible to evaluate these estimates with observed data as a reference. The literature has aimed to model deviations from CIP with a single-equation model, and this paper is no exception. Modelling distinct covered FX markets may require further econometric tools, such as a time-varying multivariate model.

Market participants can use the results presented here to estimate the minimum profits they require to obtain in carry trade, or the maximum losses they are willing to underwrite. During periods of financial stress, for example, the estimated neutral band will widen in accordance to a diminished risk appetite. Deviations from CIP can still be used to price assets like covered funding, and compute the markup required using the neutral band. Low stress, would mean a low markup, high stress would demand a high markup.

Financial authorities can use the results as a surveillance tool for the appropriate functioning of markets for the assets involved in covered interest parity. Policymakers requiring to monitor on a high frequency basis liquidity conditions on the FX or covered FX market can consider deviations from CIP outside the neutral band as signal of stress. Thus, the neutral band can assist in making a decision about FX market interventions or other revisions in liquidity arrangements.

The results presented here also provide some insights for open macroeconomic modeling. The equilibrium conditions used as a benchmark for calibration and simulation of models, can be relaxed to include changes in variance. While this is is a tall order task, given that many quantitative general equilibrium models are based on linearising equilibrium conditions, it can payoff in terms of flexibility to explicitly model shocks and structural changes.

Future research can build upon these results and continue exploring the possible uses of the neutral band. There is work to be done to consider a number of alternative empirical models and data sets. These, in turn, may allow conditioning on the idiosyncratic features of the analysed currency.

References

Ardia, D. (2008). *Financial Risk Management with Bayesian Estimation of GARCH Models: Theory and Applications*. Lecture Notes in Economics and Mathematical Systems 612. Springer-Verlag.

Avdjiev, S., W. Du, C. Koch, and H. S. Shin (2019). The dollar, bank leverage, and deviations from covered interest parity. *American Economic Review: Insights 1*(2), 193–208.

BIS (2019). Mgn: Margin requirements 15 dec 2019. Technical report, Bank for International Settlements. Basel Committee on Banking Supervision.

BIS (2022). Triennial central bank survey: Foreign exchange turnover in december 2022. Technical report, Bank for International Settlements.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*(3), 307–327.

Branson, W. H. (1969). The minimum covered interest differential needed for international arbitrage activity. *Journal of Political Economy 77*(6), 1028–1035.

Brunnermeier, M. K. and L. H. Pedersen (2008, 11). Market Liquidity and Funding Liquidity. *The Review of Financial Studies 22*(6), 2201–2238.

Bush, G. and G. López Noria (2021). Uncertainty and exchange rate volatility: Evidence from mexico. *International Review of Economics & Finance 75*, 704–722.

Campbell, J. Y. (1993). Intertemporal asset pricing without consumption data. *The American Economic Review 83*(3), 487–512.

Cendese, G., P. della Corte, and T. Wang (2021). Currency mispricing and dealer balance sheets. *The Journal of Finance 76*(6), 2763–2803.

Cerutti, E. M., M. Obstfeld, and H. Zhou (2021). Covered interest parity deviations: Macrofinancial determinants. *Journal of International Economics 130*, 103447.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review 39*(4), 841–862.

Clinton, K. (1988). Transactions costs and covered interest arbitrage: theory and evidence. *Journal of Political Economy 96*(2), 358–370.

Dacco, R. and S. Satchell (1999). Why do Regime Switching Models Forecast so Badly? *Journal of Forecasting 18*, 1–16.

Deardorff, A. V. (1979). One-way arbitrage and its implications for the foreign exchange markets. *Journal of Political Economy 87*(2), 351–364.

Du, W., B. Hébert, and A. W. Huber (2022, 08). Are Intermediary Constraints Priced? *The Review of Financial Studies 36*(4), 1464–1507.

Du, W., J. Im, and J. Schreger (2018). The u.s. treasury premium. *Journal of International Economics 112*, 167–181.

Du, W. and J. Schreger (2016). Local currency sovereign risk. *The Journal of Finance 71*(3), 1027–1070.

Du, W. and J. Schreger (2022). 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*, Volume 6 of *Handbook of International Economics*, pp. 147–197. Elsevier.

Du, W., A. Tepper, and A. Verdelhan (2018). Deviations from covered interest rate parity. *The Journal of Finance 73*(3), 915–957.

Einzig, P. (1967). *A dynamic theory of forward exchange*. Macmillan.

Elliott, G. and A. Timmermann (2016). *Economic Forecasting*. Princeton University Press.

Epstein, L. G. and S. E. Zin (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica 57*(4), 937–969.

Fratzscher, M., O. Gloede, L. Menkhoff, L. Sarno, and T. Stöhr (2019, January). When is foreign exchange intervention effective? evidence from 33 countries. *American Economic Journal: Macroeconomics 11*(1), 132–56.

Frenkel, J. A. (1973). Elasticities and the interest parity theory. *Journal of Political Economy 81*(3), 741–747.

Frenkel, J. A. and R. M. Levich (1975). Covered interest arbitrage: Unexploited profits?. *Journal of Political Economy 83*(2), 325 – 338.

Frenkel, J. A. and R. M. Levich (1977). Transaction costs and interest arbitrage: Tranquil versus turbulent periods. *Journal of Political Economy 85*(6), 1209 – 1226.

Geweke, J. and G. Amisano (2010). Comparing and evaluating bayesian predictive distributions of asset returns. *International Journal of Forecasting 26*(2), 216 – 230. Special Issue: Bayesian Forecasting in Economics.

Granger, C. and T. Teräsvirta (1993). *Modelling Nonlinear Economic Relationships*. Oxford University Press.

He, Z. and A. Krishnamurthy (2018). Intermediary asset pricing and the financial crisis. *Annual Review of Financial Economics 10*(Volume 10, 2018), 173–197.

Hernandez, J. R. (2014). Peso-dollar forward market analysis: Explaining arbitrage opportunities during the financial crisis. *Banco de México Working Papers* (No. 2014-09).

Hosszejni, D. and G. Kastner (2021). Modeling univariate and multivariate stochastic volatility in r with stochvol and factorstochvol. *Journal of Statistical Software 100*(12), 1–34.

IMF (2020). United states: Financial system stability assesment. Technical report, International Monetary Fund.

Juhl, T., W. Miles, and M. D. Weidenmier (2006). Covered interest arbitrage: then versus now. *Economica 73*(290), 341–352.

Keynes, J. M. (1923). *A Tract on Monetary Reform*. Macmillan.

Kim, S., N. Shephard, and S. Chib (1998). Stochastic volatility: Likelihood inference and comparison with arch models. *The Review of Economic Studies 65*(3), 361–393.

Levich, R. M. (2017). Cip: Then and now, a brief survey of measuring and exploiting deviations from covered interest parity. Technical report, Remarks prepared for the BIS Symposium: CIP - RIP?

Omori, Y., S. Chib, N. Shephard, and J. Nakajima (2007). Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics 140*(2), 425–449.

Peel, D. A. and M. P. Taylor (2002). Covered interest rate arbitrage in the interwar period and the keynes-einzig conjecture. *Journal of Money, Credit & Banking (Ohio State University Press) 34*(1), 51 – 75.

Shephard, N. (2015). Martingale unobserved component models. In S. J. Koopman and N. Shephard (Eds.), *Unobserved Components and Time Series Econometrics*, pp. 218–249. Oxford: Oxford University Press.

Stigler, M. (2019). Nonlinear time series in r: Threshold cointegration with tsdyn. In H. D. Vinod and C. Rao (Eds.), *Handbook of Statistics, Volume 41*, Volume 42, pp. 229–264. Elsevier.

Tong, H. (1990). *Nonlinear Time Series: A Dynamical System Approach*. Claredon Press.

Appendix

A Summary Statistics and Estimates for the MXN-USD

	Sample size	Mean	S.D.	Max	75-q	25-q	Min
GBP-USD	5,523	-13.54	23.61	44.09	0.72	-22.19	-280.81
MXN-USD	4,785	-37.38	67.44	347.47	2.59	-74.23	-74.23

Table 6: Summary statistics for the cross-currency basis x_t. Al statistics are in basis points. S.D. is the standard deviation. 75-q and 25-q refer to the 75th and 25th quantile.

		TAR regimes					
		Lower $\kappa^D = -24.15$	Mid	Upper $\kappa^D = 4.73$	GARCH	SV	SVL
β							
	β_0	-8.022	-0.098	0.265	73.399	-0.419	-0.464
		(1.612)	(1.397)	(0.298)	(14.703)	(0.126)	(0.130)
	β_1	0.933	0.9943	0.949	0.999	0.990	0.989
		(0.012)	(0.023)	(0.006)	(0.001)	0.002	(0.002)
θ							
	θ_0	13.881	13.881	13.881	1.222	3.790	-9.590
					(0.638)	(0.187)	(1.050)
	θ_1				0.114	0.974	1.000
					(0.031)	(0.005)	(0.000)
	θ_2				0.883	-	-
					(0.033)	-	-
	σ_η					0.084	0.056
						(0.012)	0.008
	ρ						-0.085
							(0.049)

Table 7: Estimates for the covered MXN-USD (3.5) and (3.6). Standard errors in parenthesis. Details on the estimation are as in Table 7 are contained in the Appendix.

Model	Empirical Coverage	LR_{uc}	LR_{ind}	LR_{cc}
TAR	97.08	47.324	220.365	267.748
GARCH	95.55	2.900	8.361	11.353
SV	94.82	0.283	5.266	5.655
SVL	94.57	1.629	4.701	6.443
Critical values (95%)		3.841	3.841	5.991

Table 8: Coverage ratio tests and empirical coverage for the covered MXN-USD. The empirical coverage of each model is defined as the number of observations that fall within the neutral band. Christoffersen (1998) tests for the congruence of the theoretical 95% and the percentage of forecast errors, LR_{uc}, the test for independence of the forecast errors, LR_{ind}, and the test for the correct coverage conditional om the forecasts error being independent LR_{cc}. The null hypothesis is correct coverage. Critical values for the 95% are obtained from the χ^2_r distribution with r degrees of freedom.

Model	Q_{LS}	Accuracy	*t-stat*	p-value
		GARCH vs TAR	0.543	4.867×10^{-15}
TAR	-4.131	SV vs TAR	0.594	5.513×10^{-16}
GARCH	-3.588	SVL vs TAR	0.593	5.814×10^{-16}
SV	-3.536	SV vs GARCH	0.051	3.693×10^{-4}
SVL	-3.537	SVL vs GARCH	0.050	4.153×10^{-4}
		SV vs SVL	0.001	0.097

Table 9: Log-score comparison for the covered MXN-USD. The Q_{LS} statistic (4.2) summarises the one-step-ahead forecast precision of each model. Large precision corresponds to large Q_{LS}. The *t-stat* is the Diebold-Mariano test statistic that compares if two models have the same Q_{LS} statistically. The null hypothesis is: two models have the same Q_{LS}. Heteroskedasticity-Autocorrelation robust standard errors are used in the comparison.

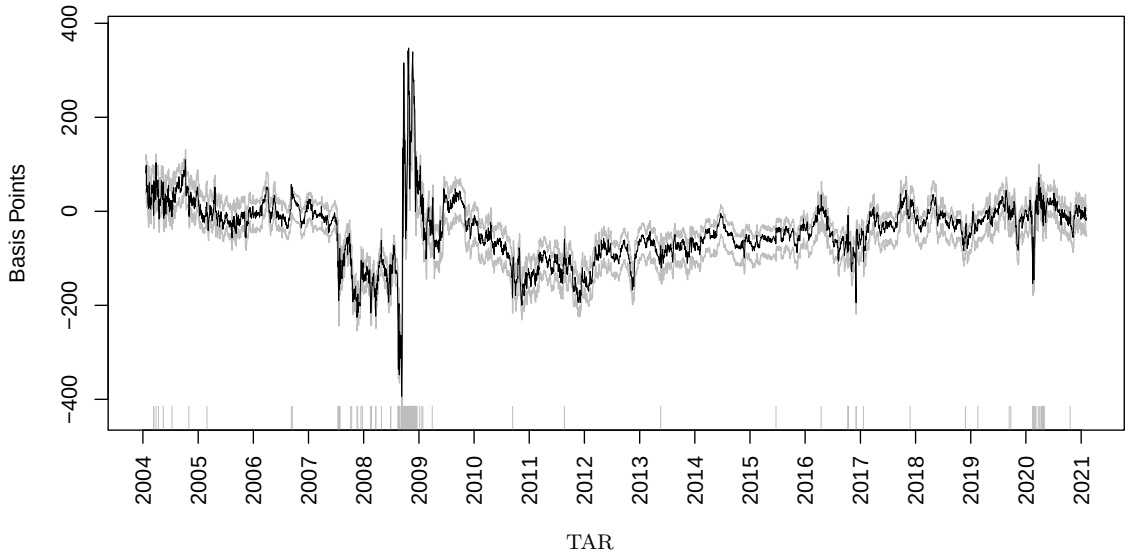

Figure 8: Black: Deviations from CIP, x_t, in basis points. Grey: TAR model Neutral band estimates for MXN-USD, defined as the 97.5% and 2.5% quantile estimate of the one step-ahead predictive distribution $\hat{p}(x_t|\mathcal{F}_{t-1})$. The grey vertical markers are occurrences of $x_t > \hat{\bar{b}}_t$ or $x_t < \hat{\underline{b}}_t$ in basis points. Source: Own estimates and Du and Schreger (2016) and Du et al. (2018).

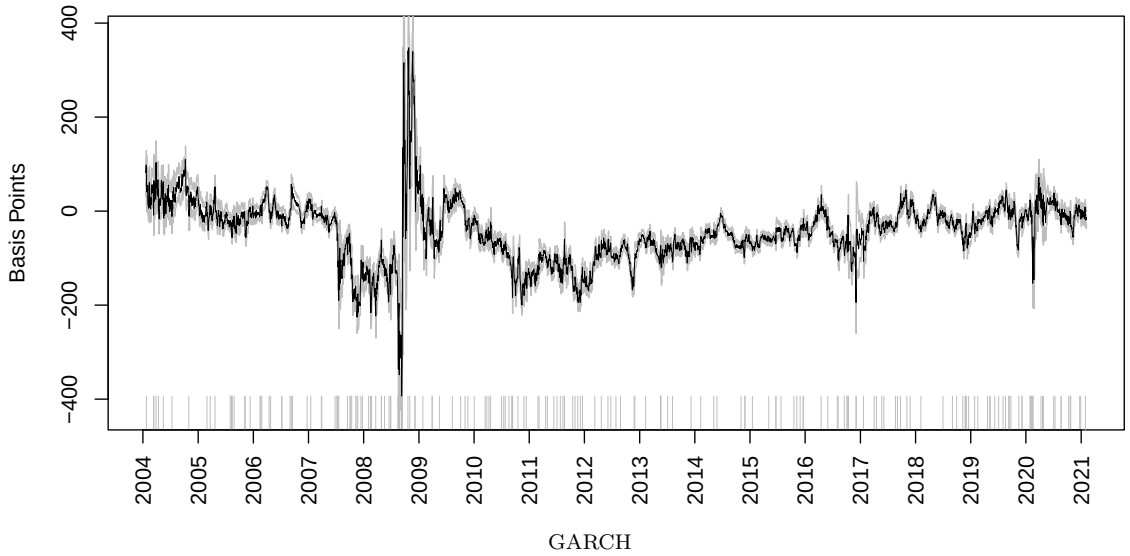

Figure 9: Black: Deviations from CIP, x_t, in basis points. Grey: GARCH model Neutral band estimates for MXN-USD, defined as the 97.5% and 2.5% quantile estimate of the one step-ahead predictive distribution $\hat{p}(x_t|\mathcal{F}_{t-1})$. The grey vertical markers are occurrences of $x_t > \hat{\bar{b}}_t$ or $x_t < \hat{\underline{b}}_t$ in basis points. Source: Own estimates and Du and Schreger (2016) and Du et al. (2018).

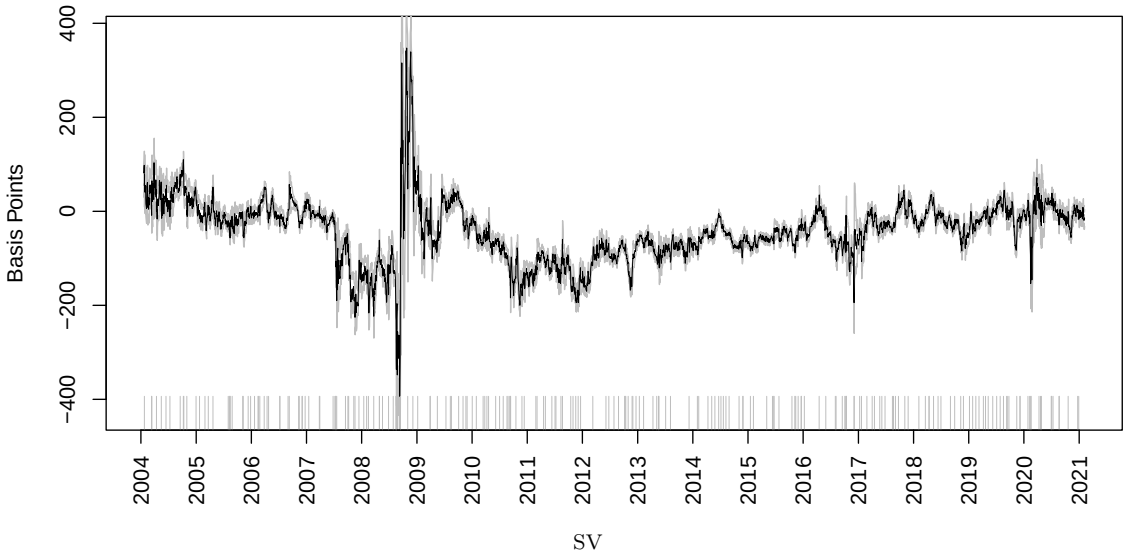

Figure 10: Black: Deviations from CIP, x_t, in basis points. Grey: SV model Neutral band estimates for MXN-USD, defined as the 97.5% and 2.5% quantile estimate of the one step-ahead predictive distribution $\hat{p}(x_t|\mathcal{F}_{t-1})$. The grey vertical markers are occurrences of $x_t > \hat{\bar{b}}_t$ or $x_t < \hat{\underline{b}}_t$ in basis points. Source: Own estimates and Du and Schreger (2016) and Du et al. (2018).

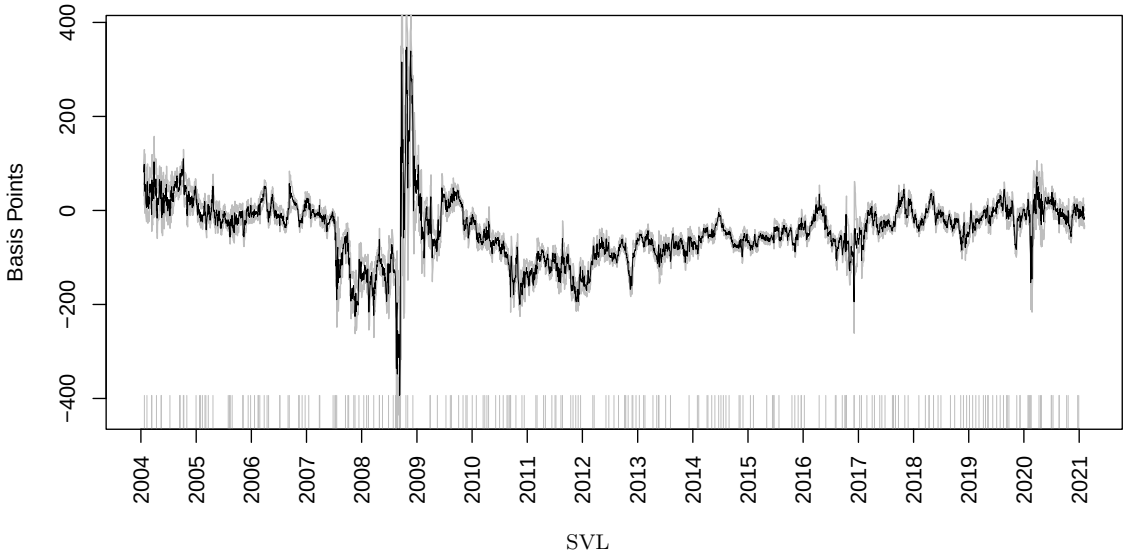

Figure 11: Black: Deviations from CIP, x_t, in basis points. Grey: SVL model Neutral band estimates for MXN-USD, defined as the 97.5% and 2.5% quantile estimate of the one step-ahead predictive distribution $\hat{p}(x_t|\mathcal{F}_{t-1})$. The grey vertical markers are occurrences of $x_t > \hat{\bar{b}}_t$ or $x_t < \hat{\underline{b}}_t$ in basis points. Source: Own estimates and Du and Schreger (2016) and Du et al. (2018).

B Appendix Details on Estimates GBP-USD

β^T	Estimate	S.E.	t-statistic	p-value
β_0^L	-2.606	0.371	-7.008	0.000
β_1^L	0.929	0.007	132.947	0.000
β_0^M	-0.551	0.171	-3.226	0.001
β_1^M	0.931	0.015	60.824	0.000
β_0^U	1.641	0.535	3.067	0.002
β_1^U	0.754	0.047	16.050	0.000
$Q(30)$	379.8			0.000
$\theta_0 = 6.811$, $\kappa^D = -24.150$, $\kappa^U = 4.735$				

Table 10: Non-linear least squares estimates for the TAR model parameters β^T and lower and upper threshold estimates κ^L, κ^U. θ_0 is the error unconditional variance. S.E. is the standard error. $Q(l)$ is the Box-Ljung autocorrelation test statistic on the residuals for l lags.

		Estimate	S.E.	t-statistic	p-value
β^G					
	β_0^G	-3.822	1.778	-2.150	0.039
	β_1^G	0.974	0.004	261.444	0.000
θ^G					
	θ_0^G	0.231	0.145	1.592	0.112
	θ_1^G	0.125	0.032	3.867	0.000
	θ_2^G	0.873	0.035	24.808	0.000
	$Q(30)$	205.44			0.000

Table 11: Maximum likelihood estimates for the GARCH model parameters, β^G, θ^G. S.E. is standard error. $Q(l)$ is the Box-Ljung autocorrelation test statistic on the residuals for l lags. The parameters are estimated by maximum likelihood through the *fGarch* R package from *Rmetrics*. The elements of β^G follow a multivariate normal distribution, while $\theta_0^G, \theta_1^G, \theta_2^G$ follow a truncated multivariate normal distribution.

		Mean	S.E.	Autocorrelation lag:	500	1000	2000
β^S							
	β_0^S	-0.090	0.040		0.000	0.001	-0.001
	β_1^S	0.979	0.003		0.000	-0.004	-0.002
θ^S				95% C.I.: $\lvert 0.0273 \rvert$			
	θ_0^S	2.230	0.193		-0.037	-0.076	0.042
	θ_1^S	0.975	0.005		-0.022	-0.030	-0.000
	σ_η^S	0.096	0.012		0.016	0.011	-0.029

Table 12: Summary statistics for full sample estimates for the SV model parameters: posterior mean of β^S, θ^S; S.E. is the Monte Carlo standard error; and autocorrelation of the MCMC chain of length 10^4. C.I. is the confidence interval of the autocorrelation function. The prior distributions are $(\sigma_\eta^S)^2 \sim \mathcal{IG}(5, 0.05)$ (inverse-gamma distribution); $\theta_1^S \sim \mathcal{B}(20, 1.5)$ (beta distribution); and $\theta_0^S \sim \mathcal{N}(0, 10)$. The prior distributions and hyperparameters are chosen to guarantee that $(\sigma_\eta^S)^2 > 0$ and $\theta_1^S \in (-1, 1)$. Draws for the autoregressive mean and the volatility process obtained with the R package *stochvol* from Hosszejni and Kastner (2021). The parameters for the conditional mean have normal prior and posterior distributions.

		Mean	S.E.	Autocorrelation lag:	500	1000	2000
β^L							
	β_0^L	-0.105	0.042		-0.007	0.001	-0.011
	β_1^L	0.978	0.003		0.019	0.012	0.004
θ^L				95% C.I.: $\lvert 0.0273 \rvert$			
	θ_0^L	2.370	0.206		-0.083	-0.075	0.014
	θ_1^L	0.987	0.003		-0.060	-0.029	-0.036
	σ_η^L	0.041	0.006		0.046	-0.096	0.005
	ρ	-0.063	0.049		0.038	-0.025	0.006

Table 13: Summary statistics for full sample estimates for the SVL model parameters: posterior mean of β^L, θ^L; S.E. is the Monte Carlo standard error; and autocorrelation of the MCMC chain of length 10^4. C.I. is the confidence interval of the autocorrelation function. The prior distribution and hyperparameters ar chosen to obtain $\rho \in (-1, 1)$, in particular $(\rho + 1)/2 \sim \mathcal{B}(4, 4)$. The prior distributions for the rest of the parameters are the same as those in the SV. Draws for volatility are estimated with the *stochvol* package (Hosszejni and Kastner, 2021).

TAR: The steps to obtain the b_t estimate and the likelihood are as follows:

1. Define the training sample $t = 1, \ldots, t_0 = 360$.
2. Obtain NLLS point estimates of $\left(\beta^{T\prime}, \theta^{T\prime}, \kappa^D, \kappa^U\right)'$ using the R package *tsDyn* provided by Stigler (2019). Here $i \in \{D, M, U\}$ as in (3.8).
3. Obtain the one-step ahead forecast $\widehat{x}_{t|t-1,\kappa}$ and its variance $\widehat{V}_{t|t-1,\kappa}$, defined in (3.3) and (3.4) for each of the three regimes.
4. Estimate $\widehat{p}\left(x_t|\mathcal{F}_{t-1}\right)$ by obtaining draws, $\widehat{x}_{t|t-1,\kappa}^{(j)}$, from $\mathcal{N}\left(\widehat{x}_{t|t-1,\kappa}, \widehat{V}_{t|t-1,\kappa}\right)$, $j = 1, \ldots, M$, with $M = 10^4$.
5. Compute the log-likelihood of the next observation being contained in the forecast distribution (log-score) as $LS_{t_0} = 1/M \sum_{j=1}^M \ln \widehat{p}\left(x_t|\widehat{x}_{t|t-1,\kappa}^{(j)}, \widehat{V}_{t|t-1,\kappa}, \mathcal{F}_{t-1}\right)$.
6. Define $\hat{\bar{b}}_{t|t-1}, \hat{\underline{b}}_{t|t-1}$ as the 0.975 and 0.025 estimated quantiles from (3.2), respectively, and $\hat{b}_{t|t-1} = \hat{\bar{b}}_{t|t-1} - \hat{\underline{b}}_{t|t-1}$.
7. Add an observation to t_0 and go back to 2.

Repeat the steps until $t = n - 1$.

GARCH: The steps to obtain the b_t estimate and the likelihood are as follows:

1. Define the training sample $t = 1, \ldots, t_0 = 360$.
2. Obtain the maximum likelihood estimates of β^G, θ^G and h^G using the R package *rugarch* provided by Stigler (2019). This is given by β^G, θ^G and h_t^G.
3. Estimate $\widehat{p}\left(x_t|\mathcal{F}_{t-1}\right)$ by obtaining $j = 1, \ldots, M$ draws, from $\mathcal{N}\left(\widehat{x}_{t|t-1}, \widehat{V}_{t|t-1}\right)$ for each j with $M = 10^4$:
   (a) Compute $\hat{h}_{t|t-1}^G = \hat{\theta}_0^G + \hat{\theta}_1^G x_{t-1}^2 + \hat{\theta}_2^G \hat{h}_{t-1|t-2}^G$.
   (b) Compute $\hat{x}_{t|t-1} = \hat{\beta}_0^G + \hat{\beta}_1^G x_{t-1} + \sqrt{\hat{h}_{t|t-1}^G}\varepsilon_t^{(j)}$ and $\varepsilon_t^{(j)}$ is a draw from $\mathcal{N}(0,1)$, defined in (3.3) and (3.4).
4. Estimate $\widehat{p}\left(x_t|\mathcal{F}_{t-1}\right)$ by obtaining draws, $\widehat{x}_{t|t-1,}^{(j)}$, from $\mathcal{N}\left(\widehat{x}_{t|t-1,}, \widehat{V}_{t|t-1,}\right)$, $j = 1, \ldots, M$, with $M = 10^4$.
5. Compute the log-likelihood of the next observation being contained in the forecast distribution (log-score) as $LS_{t_0} = 1/M \sum_{j=1}^{M} \ln \widehat{p}\left(x_t|\widehat{x}_{t|t-1,}^{(j)}, \widehat{V}_{t|t-1,}, \mathcal{F}_{t-1}\right)$.
6. Define $\hat{\bar{b}}_{t|t-1}, \hat{\underline{b}}_{t|t-1}$ as the 0.975 and 0.025 estimated quantiles from (3.2), respectively, and $\hat{b}_{t|t-1} = \hat{\bar{b}}_{t|t-1} - \hat{\underline{b}}_{t|t-1}$.
7. Add an observation to t_0 and go back to 2.

Repeat the steps until $t = n - 1$.

SV: The steps to obtain the b_t estimate and the likelihood are as follows:

1. Define the training sample $t = 1, \ldots, t_0 = 360$.
2. Obtain estimates for the posterior distributions of θ and h using the R package *stochvol* provided by Hosszejni and Kastner (2021). Note that this is given by $\left\{\theta^{S(j)}\right\}_{j=1}^{M}$ and $\left\{h_t^{S(j)}\right\}_{j=1}^{M}$ with $M = 10^4$, a burn-in sample of 10^3, and $t = 1, \ldots, t_0$. For this model h_t is the log of the conditional variance.
3. Estimate $\widehat{p}\left(x_t|\mathcal{F}_{t-1}\right)$ by obtaining a draw from $\mathcal{N}\left(x_{t|t-1}^{(j)}, V_{t|t-1}^{(j)}\right)$ for each j:
   (a) Compute $h_{t|t-1}^{S(j)} = \theta_0^{S(j)} + \theta_1^{S(j)}\left(h_{t-1|t-2}^{S(j)} - \theta_0^{S(j)}\right) + \sigma_\eta^{S(j)}\eta_t^{(j)}$, where $\eta_t^{(j)}$ is a draw from $\mathcal{N}(0,1)$.
   (b) Compute $x_{t|t-1}^{(j)} = \beta_0^{S(j)} + \beta_1^{S(j)} x_{t-1} + \sqrt{\exp\left(h_{t|t-1}^{S(j)}\right)}\varepsilon_t^{(j)}$ where $\varepsilon_t^{(j)}$ is a draw from $\mathcal{N}(0,1)$.
4. Estimate $\widehat{p}\left(x_t|\mathcal{F}_{t-1}\right)$ by obtaining draws, $\widehat{x}_{t|t-1,}^{(j)}$, from $\mathcal{N}\left(\widehat{x}_{t|t-1,}^{(j)}, \widehat{V}_{t|t-1,}^{(j)}\right)$, $j = 1, \ldots, M$, with $M = 10^4$ and mean and variance defined in (3.3) and (3.4).
5. Compute the log-likelihood of the next observation being contained in the forecast distribution (log-score) as $LS_{t_0} = 1/M \sum_{j=1}^{M} \ln \widehat{p}\left(x_t|\widehat{x}_{t|t-1,}^{(j)}, \widehat{V}_{t|t-1,}^{(j)}, \mathcal{F}_{t-1}\right)$.
6. Define $\hat{\bar{b}}_{t|t-1}, \hat{\underline{b}}_{t|t-1}$ as the 0.975 and 0.025 estimated quantiles from (3.2), respectively, and $\hat{b}_{t|t-1} = \hat{\bar{b}}_{t|t-1} - \hat{\underline{b}}_{t|t-1}$.
7. Add an observation to t_0 and go back to 2.

Repeat the steps until $t = n - 1$.

SVL: The steps to obtain the b_t estimate and the likelihood are as follows:

1. Define the training sample $t = 1, \ldots, t_0 = 360$.
2. Obtain estimates for the posterior distributions of θ and h using the R package *stochvol* provided by Hosszejni and Kastner (2021). Note that this is given by $\left\{\theta^{L(j)}\right\}_{j=1}^{M}$ and $\left\{h_t^{L(j)}\right\}_{j=1}^{M}$ with $M = 10^4$, a burn-in sample of 10^3, and $t = 1, \ldots, t_0$. For this model h_t is the log of the conditional variance.
3. Estimate $\widehat{p}\left(x_t|\mathcal{F}_{t-1}\right)$ by obtaining a draw from $\mathcal{N}\left(x_{t|t-1}^{(j)}, V_{t|t-1}^{(j)}\right)$ for each j:
   (a) Draw $\left(\eta_t^{(j)}, \varepsilon_t^{(j)}\right)'$ from a bi-variate normal distribution, where elements display correlation ρ at each period t.
   (b) Compute $h_{t|t-1}^{S(j)} = \theta_0^{S(j)} + \theta_1^{S(j)}\left(h_{t-1|t-2}^{S(j)} - \theta_0^{S(j)}\right) + \sigma_\eta^{S(j)}\eta_t^{(j)}$.
   (c) Compute $x_{t|t-1}^{(j)} = \beta_0^{S(j)} + \beta_1^{S(j)} x_{t-1} + \sqrt{\exp\left(h_{t|t-1}^{S(j)}\right)}\varepsilon_t^{(j)}$.
4. Estimate $\widehat{p}\left(x_t|\mathcal{F}_{t-1}\right)$ by obtaining draws, $\widehat{x}_{t|t-1,}^{(j)}$, from $\mathcal{N}\left(\widehat{x}_{t|t-1,}^{(j)}, \widehat{V}_{t|t-1,}^{(j)}\right)$, $j = 1, \ldots, M$, with $M = 10^4$ and mean and variance defined in (3.3) and (3.4).
5. Compute the log-likelihood of the next observation being contained in the forecast distribution (log-score) as $LS_{t_0} = 1/M \sum_{j=1}^{M} \ln \widehat{p}\left(x_t|\widehat{x}_{t|t-1,}^{(j)}, \widehat{V}_{t|t-1,}^{(j)}, \mathcal{F}_{t-1}\right)$.
6. Define $\hat{\bar{b}}_{t|t-1}, \hat{\underline{b}}_{t|t-1}$ as the 0.975 and 0.025 estimated quantiles from (3.2), respectively, and $\hat{b}_{t|t-1} = \hat{\bar{b}}_{t|t-1} - \hat{\underline{b}}_{t|t-1}$.
7. Add an observation to t_0 and go back to 2.

Repeat the steps until $t = n - 1$.

Previous volumes in this series

1205 August 2024	The Measure Matters: Differences in the Passthrough of Inflation Expectations in Colombia	Andres Sanchez-Jabba and Erick Villabon-Hinestroza
1204 August 2024	Climate Policies, Labor Markets, and Macroeconomic Outcomes in Emerging Economies	Alan Finkelstein Shapiro and Victoria Nuguer
1203 August 2024	Strike while the Iron is Hot: Optimal Monetary Policy with a Nonlinear Phillips Curve	Peter Karadi, Anton Nakov, Galo Nuno, Ernesto Pasten, and Dominik Thaler
1202 August 2024	Are low interest rates firing back? Interest rate risk in the banking book and bank lending in a rising interest rate environment	Lara Coulier, Cosimo Pancaro and Alessio Reghezza
1201 July 2024	Crypto Exchange Tokens	Rodney Garratt, Maarten R.C. van Oordt
1200 July 2024	Financial inclusion transitions in Peru: does labor informality play a role?	Jose Aurazo and Farid Gasmi
1199 July 2024	New spare tires: local currency credit as a global shock absorber	Stefan Avdjiev, John Burger and Bryan Hardy
1198 July 2024	Sovereign green bonds: a catalyst for sustainable debt market development?	Gong Cheng, Torsten Ehlers, Frank Packer and Yanzhe Xiao
1197 July 2024	The gen AI gender gap	Iñaki Aldasoro, Olivier Armantier, Sebastian Doerr, Leonardo Gambacorta and Tommaso Oliviero
1196 July 2024	Digital payments, informality and economic growth	Ana Aguilar, Jon Frost, Rafael Guerra, Steven Kamin and Alexandre Tombini
1195 July 2024	The asymmetric and persistent effects of Fed policy on global bond yields	Tobias Adrian, Gaston Gelos, Nora Lamersdorf, Emanuel Moench
1194 June 2024	Intelligent financial system: how AI is transforming finance	Iñaki Aldasoro, Leonardo Gambacorta, Anton Korinek, Vatsala Shreeti and Merlin Stein
1193 June 2024	Aging gracefully: steering the banking sector through demographic shifts	Christian Schmieder and Patrick A Imam
1192 June 2024	Sectoral heterogeneity in the wage-price pass-through: Evidence from Euro area	Miguel Ampudia, Marco Lombardi and Théodore Renault

All volumes are available on our website www.bis.org.