Covered Interest Parity: A Stochastic Volatility Approach to Estimate the Neutral Band

Hernández, Juan R.

Banco de Mexico

2020

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

Juan R. Hernández
Banco de México

November 2019

Abstract

The neutral band is the interval where deviations from Covered Interest Parity (CIP) are not considered meaningful arbitrage opportunities. The band is determined by transaction costs and risk associated to arbitrage. Seemingly large deviations from CIP in the foreign exchange markets for the US Dollar crosses with Sterling, Euro and Mexican Peso have been the norm ever since the onset of the Global Financial Crisis. This topic has attracted a great deal of attention in the literature. There are no estimates of the neutral band to assess whether deviations from CIP reflect actual arbitrage opportunities, however. This paper proposes an estimate of the neutral band based on the one-step-ahead density forecast obtained from a single equation stochastic volatility model. Comparison across econometric models is made using the log-score statistic and the probability integral transformation, computed from the predictive likelihood function. The stochastic volatility models have the best fit and forecasting performance, hence superior neutral band estimates.

Keywords: Covered interest parity; stochastic volatility; forward filtering backward smoothing; auxiliary particle filter; density forecast.

J.E.L. classification number. C53, C58, F31, F37.

Acknowledgements: Special thanks to Nicolás Amoroso, Alfonso Cebreros, Raúl Ibarra, Mauricio Olivares and Vladimir Rodríguez for helpful discussion and comments on previous drafts. Comments from two internal referees, participants at ITAM's Seminario Aleatorio, and the 4th Market Microstuctre-Nonlinear Dynamics Workshop are gratefully acknowledged. Andrea Miranda, Ezequiel Piedras, Javier Pérez and María Diego-Fernández provided excellent research assistance. All remaining errors are my own.

Address for Correspondence: Juan R. Hernández, Banco de México, Dirección General de Investigación Económica, Av. 5 de Mayo 18, Colonia Centro. Mexico City, Mexico.
E-mail: juan.hernandez@banxico.org.mx

*Opinions contained in this paper correspond solely to the author and do not necessarily reflect the point of view of Banco de México.

1 Introduction

Up to the outset of the Global Financial Crisis in the summer of 2007, the covered interest rate parity (CIP) was among the most reliable relations in international finance. Events that unfolded subsequently seem to have lasting effects on said relation. Significant measured deviations from CIP in the foreign exchange (FX) markets for "reserve" currencies crosses with the US Dollar (USD) have been the norm, rather than the exception. The CIP is a cornerstone of international finance, a bellwether of market efficiency, and is used, for example, to price forward contracts, hence, informing on liquidity conditions prevailing in the FX market. As outlined by Levich (2017), deviations from CIP have motivated a growing literature aiming to answer questions such as: (i) What happened to CIP? (ii) What is behind the larger disparity in CIP? (iii) What does it mean for pricing of assets that rely on the CIP being satisfied?

Levich (2017) also posed the following, perhaps subtler, question: "Are the CIP deviations we see today only 'measured deviations' bounded within a neutral band that captures all the costs and risks of arbitrage?" This paper proposes to answer this question by estimating the size of the neutral band, thus complementing the extant literature in addressing questions (i) to (iii). It has been long acknowledged that deviations from CIP may be present while satisfying the absence of arbitrage opportunities, as long as they are contained within the neutral band. The width of said band is, in turn, determined by a number of factors, such as transaction costs, risk aversion, and uncertainty. This paper estimates the width of the neutral band for three FX markets with a high volume of transactions according to BIS (2016): Two markets that relate "reserve" currencies, Sterling (GBP-USD) and Euro (EUR-USD), and one market from an Emerging Market Economy, Mexican Peso (MXN-USD). These FX markets have displayed persistent deviations from CIP in the years 2008-2017.

Literature on estimating the band around deviations from CIP has focused on different periods of the 20th century. Seemingly absent from this literature, however, is a measure of the width of the corresponding neutral band for the *post* Global Financial Crisis period. These studies have focused on estimating the band around deviations from CIP using either of two approaches. First, the "counting" approach which entails declaring the upper- and lower-limits of the neutral band as the levels of deviations such that 95% of the measured deviations are contained. The second approach is based on econometric estimates of the values at which a change of regime in or out the neutral band is determined. This paper synthesises the two approaches by estimating the neutral band as the one-step-ahead 95% density forecast for deviations from CIP.

The main difference between this paper and previous work estimating the neutral band, most notably Peel and Taylor (2002), is the following. Instead of testing if deviations from CIP are within given fixed estimates of the upper- and lower-bounds of the neutral band, the focus here is on determining if the majority of observed deviations from CIP are replicable by a model in which arbitrage is absent. That is, if deviations from CIP are a Martingale process.

Every found paper focusing on the post-Global Financial Crisis period can be motivated by questions (i)-(iii) listed above. None, however, embarks on estimating whether deviations from CIP are bounded within the neutral band that captures all costs and risks of arbitrage. The main contribution of this paper is the estimate of the neutral band for the period 2003-2017 for the USD-GBP, USD-EUR, and USD-MXN pairs. Estimation is carried out with interest rates for interbank debt for each currency pair.

The best model to estimate the neutral band, for each currency pair, is obtained from the comparison between the baseline threshold autoregressive model and (i) a local-level model with a time-varying mean (estimated through maximum likelihood and through Markov Chain Monte Carlo); and (ii) a stochastic volatility model (where volatility is estimated by forward-filtering-backward-smoothing and auxiliary particle filtering). The time-varying approach to estimation aims to model that not all deviations from CIP are caused by the same factors as acknowledged by Levich (2017). The goodness of fit of each model is assessed through the log-score and the probability integral transformation of the one-step-ahead density forecast, estimated in an expanding window sample.

Estimation of the neutral band as the one-step-ahead density forecast synthesises both the "counting" and econometric approaches previously followed. The former is embedded by taking the 95% density forecast, while the latter is embedded in the econometric methods applied to obtain the density forecast. Risk, hedging demand, and changes in the institutional framework may be reflected in a rise of uncertainty, which, in turn, is a determinant of the width of the neutral band. Hence, by modelling explicitly the mean and volatility processes underlying the CIP, a reliable estimate of the neutral band is obtained.

This estimation strategy has several advantages and desirable features. First, both the local-level and the stochastic volatility models are representations of a Martingale process (Shephard, 2013). In turn, it has been long-established that no-arbitrage is equivalent to Martingale pricing (Duffie, 2010). Thus, if the models fit the observed data, there should be no arbitrage and the CIP should be satisfied. Second, by identifying those observations that violate the Martingale property, a narrower focus on explaining deviations from CIP may be undertaken in future work. The extant literature aims to explain these deviations (possibly a discrete-process) with continuous omitted variables. Third, estimating and forecasting with an expanding window, replicates the way market participants assess time-varying transaction costs and risk which are permanently changing. Finally, the models are suited to deal with financial data collected at high frequency.

Results show that the stochastic volatility models fit the data better, both in- and out-of sample. Hence, these models produce best neutral band estimates according to the log-score statistic and the probability integral transformation computed from the predictive likelihood functions. Unlike the local level and threshold autoregressive models, the neutral band estimates increase in periods of domestic or global financial stress. Moreover, the stochastic volatility model is able to replicate that arbitrage opportunities are short-lived since devi-

ations from CIP outside the neutral band estimate are scattered through the sample. In contrast, the local level and threshold autoregressive models yield a neutral band which implies long-lived (clustered) arbitrage opportunities. The neutral band implied by these models is implausibly large, since the estimated levels during and post-Global Financial Crisis are remarkably similar.

The paper is organised as follows. The rest of this section contains a brief literature review. In Section 2 theoretical foundations of the CIP are briefly reviewed and there is also discussion on how the neutral band around deviations from CIP treated in this paper, along with the data, are presented. The econometric analysis is contained in Section 3, wherein a detailed explanation of the methods used to estimate the neutral band is provided, along with the results from estimation and both the log-score and probability integral transformation evaluation of the density forecast is provided. Section 4 presents some concluding remarks.

The following notation conventions are used extensively in the rest of the paper: $x = \{x_t\}_{t=1}^{n}$ is a sequence of random variables with n elements; an element of the sequences is always indexed by time as x_t. An estimate of x_t conditional on the information available at period $t-1$, $\mathcal{F}_{t-1}$, is written as $x_{t|t-1}$.

Literature Review

A comprehensive literature review on the CIP and estimation of the neutral band deserves a great amount of detail and space, hence, it is beyond the scope of this paper. Such a review may be found in Levich (2017) and Claessens and Kose (2018). Closely related work to this paper, however, is briefly surveyed. Peel and Taylor (2002) and Levich (2017) trace the first efforts to estimate the neutral band back to Keynes (1923). Theory on the existence of the neutral band may be found in Einzig (1967), Branson (1969), Frenkel (1973) who labelled it, and Deardorff (1979). Estimations for the GBP-USD based on the "counting" approach described above were carried out by Frenkel and Levich (1975, 1977); Taylor (1987); and Clinton (1988).

Frenkel and Levich (1975, 1977) analyse the GBP-USD during the period 1962-1975 finding a neutral band within a 0.126 - 1.03 per cent per annum range. They estimate transaction costs associated to exploit deviations from CIP. In particular, they measure the costs of arbitrage as the sum of costs in the foreign exchange and costs in the market for securities. The former are measured through the neutral band by establishing the upper and lower limits as those in which 95% of the deviations from triangular arbitrage are contained. The latter are measured by the bid-ask spread and the brokerage fee. They conclude that data are consistent with the CIP, even when the analysis is made across exchange rate regimes. Indeed, they also conclude that it might be preferred to classify periods in terms of financial turbulence, rather than exchange rate regime. To obtain transaction costs from triangular arbitrage, an underlying assumption is that the cost structure remains constant in each regime. This assumption is relaxed in the present paper.

Taylor (1987) analyses the pairs USD-GBP, USD-Deutsche Mark, and GBP-Deutsche Mark in daily data collected 11-13 November, 1985. Taylor uses high frequency, contemporaneously sampled, quotes of interest rates and exchange rates, both spot and forward. He then recurs to a "counting" rule to determine if arbitrage opportunities existed in the sample. Findings point towards the existence of profitable opportunities in the 0-8 per cent of observations, contingent on the exchange rate and the maturity of the security used. This approach for estimating the neutral band is synthesised within the econometric approach in this paper.

Clinton (1988) analyses the exchange rates relating USD-GBP, USD-Canadian Dollar, USD-Deutsche Mark, USD-French Franc, and USD-Japanese Yen in the period November 1985-May 1986. Clinton sets out to analyse the costs of foreign exchange swaps, not considered in previous work. Clinton argues that these are the only relevant costs, since foreign exchange swaps are key determinants in pricing forward contracts. His estimates are based on the minimum of two possible bounds. The first is the one-way arbitrage advocated by Deardorff (1979), while the second is given by round-trip covered arbitrage. He estimates percentile boundaries of the neutral band in the same way as did Frankel and Levich, finding that a majority of deviations from CIP are within the 0.02-0.15 per cent per annum range. The present paper includes indirectly the role of the FX swaps, since arbitrage opportunities are in effect dislocations between a synthetic swap and the interest rate differential.

The earliest found estimation of the neutral band based on econometric techniques is Branson (1969) for the currency pairs USD-GBP and USD-Canadian Dollar for the period 1954-1964. In both cases, the estimated neutral band is 0.18 per cent per annum. Using ordinary least squares, Branson estimates the average deviations from CIP. He then labels said estimate as the average cost of transaction that an arbitrageur incurs when entering the transactions required to take advantage of an arbitrage opportunity. A key difference with respect to the present paper is the use of time-varying econometric methods to obtain neutral band estimates.

Peel and Taylor (2002) estimate the neutral band by testing econometrically the validity of the Keynes-Einzig Conjecture, which may be stated as a two-part proposition: (i) Deviations from CIP in the USD-GBP market between 1922 and 1925 will trigger arbitrage, only if the expected realised gain is larger than 1/2 per cent per annum. (ii) Deviations from CIP will dissipate only gradually since supply of funds is not perfectly elastic. The empirical analysis is carried through a single-equation estimate of the neutral band obtained with a threshold autoregressive model (TAR). This model allows to estimate three regimes for deviations from CIP. The upper and lower regimes satisfy the proposition since the estimation suggests the process is mean reverting outside a threshold of 1/2 per cent per annum. The mid regime satisfies the no-arbitrage condition since its behaviour resembles that of a Martingale process. A cointegrating multivariate TAR confirms these findings. Unlike the analysis presented in this paper, TAR estimates a neutral band that is fixed in time.

Juhl et al. (2006) use a TAR to analyse the USD-GBP pair for the Gold Standard period in 1880-1914, much in the spirit of Peel and Taylor (2002), to estimate the transaction costs which prevented arbitrage from taking place. They find evidence of a threshold for the periods 1880-1914 and 1897-1914 in the range 0.748 - 0.803 per cent per annum, whereas no threshold is found in the period 1880-1896.

As previewed above, and unlike this paper, literature analysing the CIP after the Global Financial Crisis has focused on explaining the recent behaviour of the CIP and finding possible causes of seemingly large observed deviations from the parity. That is, aiming to find an omitted variable. In the context of "reserve" currencies, the main candidate explanations are an increase in counterparty risk (Baba and Packer, 2009; Ivashina et al., 2015; Liao, 2016); liquidity constraints (Mancini-Griffoli and Ranaldo, 2011; Levich, 2012; Avdjiev et al., 2017); changes in attitudes towards risk reflected in changes in demand for hedging assets (Borio et al., 2016; Sushko et al., 2016; Cenedese et al., 2017); and changes in regulation (Du et al., 2018).

With respect to deviations from CIP observed in Emerging Markets, Mexico in particular, on the one hand Carstens (1985) and Khor and Rojas-Suarez (1991) study dislocations between long-run equilibrium conditions and observed exchange rate in a peg exchange rate regime. Analysis of the post Global Financial Crisis in Hernandez (2014) and Bush (2019), on the other hand, suggests that liquidity constraints and risk are among the causes of observed deviations from CIP. These papers do not provide estimates of the neutral band for Mexico, however.

2 Covered Interest Parity

This section formally introduces the CIP and describes the transactions needed to arbitrage away any deviations from it. In particular, two forms of arbitrage are defined, *round-trip arbitrage* as labelled by Levich (2017) and described typically in international finance literature, and *one-way arbitrage* according to Deardorff (1979). Definitions and notation that allow the modelling of the neutral band around deviations from CIP, follow. The section concludes with a description of the data used in the empirical analysis.

2.1 Risk-less Arbitrage

Assume momentarily that transaction costs are zero, there is no risk, and supply of funds for market participants is perfectly elastic. Let S_t be the spot exchange rate and let $F_{t\to k}$ be the forward rate agreed at period t with maturity of k years. Also, let $i_{t\to k}$ and $i^*_{t\to k}$ be the interest rate on a (zero-coupon) bond with maturity of k years denominated in USD and a foreign currency, respectively. Throughout the paper both S_t and $F_{t\to k}$ are measured as USD per 1 unit of the foreign currency. Assuming individuals can always borrow in USD in order

to invest in foreign bonds and the absence of financial frictions and costs, the CIP predicts

$$\frac{F_{t\to k}}{S_t} = \frac{1+i_{t\to k}}{1+i^*_{t\to k}}. \tag{2.1}$$

There are at least two sets of market participants that may exploit deviations in (2.1), (i) those in search of a profit, thus engaging in *round-trip arbitrage*; and (ii) those with future commitments denominated in foreign currency who aim to minimise the costs of the currency conversion, thus engaging in *one-way arbitrage*. Note that participants in one-way arbitrage will enter transactions related to the FX spot or future market with certainty, whereas transactions to exploit round-trip arbitrage opportunities are only engaged if profits are large enough to compensate any transaction costs or risks. Details on each form of arbitrage are given below.

Round-trip Arbitrage

In round-trip arbitrage, a market participant holds the same currency before and after a sequence of transactions. For example, assume S_t and $F_{t\to k}$ are measured in USD per 1 GBP. Assume further that

$$\frac{F_{t\to k}}{S_t}(1+i^*_{t\to k}) > 1+i_{t\to k}. \tag{2.2}$$

An investor borrowing B USD for k years at a rate $i_{t\to k}$, can buy B/S_t GBP. Then, B/S_t GBP are lent at an interest rate $i^*_{t\to k}$. Simultaneously, the investor enters a forward contract promising to deliver $(1+i^*_{t\to k})B/S_t$ GBP after k years in exchange for $F_{t\to k}B(1+i^*_{t\to k})/S_t$ USD. As a result, the investor holds $F_{t\to k}B(1+i^*_{t\to k})/S_t$ USD and he has promised to pay back $B(1+i_{t\to k})$. Since (2.2) holds, the investor will have made a (risk-less) profit: $B\left(F_{t\to k}(1+i^*_{t\to k})/S_t - (1+i_{t\to k})\right)$ USD.

One-way Arbitrage

In one-way arbitrage, a market participant will need a known amount of foreign currency at a future date, hence, holding one currency before entering arbitrage transactions and a second one after. An example is now provided. An investor holds USD knowing that he will have to deliver C^* GBP in k years; he aims to minimise the amount C of USD invested to obtain C^*. He then faces two alternatives represented by

$$\frac{C_1}{S_t}(1+i^*_{t\to k}) = C^*, \text{ and} \tag{2.3}$$

$$\frac{C_2}{F_{t\to k}}(1+i_{t\to k}) = C^*, \tag{2.4}$$

this is, the investor is looking at a cost minimisation decision $C = \min\{C_1, C_2\}$. In alternative (2.3), the investor buys C_1/S_t GBP in the spot market and invests the proceeds in a GBP security, yielding at maturity C^* GBP. In alternative (2.4), the investor buys a USD denominated security and enters a future contract in which he will receive C^* in exchange for $C_2(1+i_{t\to k})$ USD in k years. Note that, if (2.1) holds, then $C_1 = C_2$. If, however, (2.2)

holds, then $C_2 < C_1$, and the investor chooses (2.4).

Deviations from CIP and the Neutral Band

If the assumptions of zero transaction costs, no risks, and perfect elasticity of funds supply hold, theory predicts that, should a situation where (2.2) presents, for example, then (2.1) will be re-established by arbitrage. Indeed, as argued by Deardorff (1979), either form of arbitrage works to prevent permanent, or even exploding, deviations from CIP. In particular, one-way arbitrage implies that there is no-clustering of CIP deviations outside a valid neutral band.

Round-trip arbitrage implies the neutral band is not constant or even persistent, since it is only triggered if deviations are large enough to compensate for costs and risk incurred by market participants. In reality, (2.1) rarely holds exactly, in large part because the aforementioned assumptions are not satisfied. Keynes (1923) acknowledged two caveats related to the CIP proposition. First, arbitrageurs would demand a minimum profit that compensate the costs and risks inherent to the transaction. The second caveat asserts that CIP deviations would not be corrected entirely and instantaneously by transactions in the FX futures market, since supply of funds for market participants is not perfectly elastic.

To discipline the analysis, notation similar to that from Peel and Taylor (2002) is introduced. Consider the following approximation to expression (2.1)

$$P_k \frac{F_{t\to k} - S_t}{S_t} = i_{t\to k} - i^*_{t\to k},$$

where P_k adjusts to annual terms the (annualised) forward premium. Further, let *deviations* from CIP be represented by $d_{t\to k}$, and let the forward premium be represented by $\Phi_{t\to k} = P_k \frac{F_{t\to k} - S_t}{S_t}$, hence:

$$d_{t\to k} = \Phi_{t\to k} - i_{t\to k} + i^*_{t\to k}. \tag{2.5}$$

If transaction costs and risks are present, and the supply of funds for market participants is not perfectly elastic, $d_{t\to k} = 0$ may not hold exactly. Literature in international finance has long-acknowledged the latter. Indeed, previous work reviewed above declares that the CIP is satisfied as long as $d_{t\to k}$ is bounded within a *neutral band*, $w_{t\to k}$. If the observed deviation from CIP is within the neutral band (i.e. profits are smaller than the minimum required), round-trip arbitrageurs would not enter the sequence of transactions with certainty. If the observed deviation from CIP is larger than the required minimum profit, then the CIP deviation is outside the neutral band and arbitrageurs take up on the opportunity.

The magnitude of w_t, in turn, is determined by costs and risks inherent to arbitrage transactions, and availability of funds for said transactions. Formally, let $b_{U,t}$ and $b_{L,t}$ be the upper and lower limits, respectively, of $d_{t\to k}$ so that $w_{t\to k} = b_{U,t\to k} - b_{L,t\to k}$, at period t. In periods of financial stress, Levich (2017) acknowledged that w_t increases. An immediate consequence of the latter is that its values are conditional on daily available information. The w_t, however, can only be estimated from observed data, an endeavour undertaken below.

International finance literature has estimated $b_{U,t}$ and $b_{L,t}$ with either (i) a rule of thumb, whereby these values are such that 95% of observed financial deviations are contained; or (ii) non-linear econometric techniques that estimate thresholds, out of which $d_{t\to k}$ does not satisfy the Martingale property. Both techniques may be seen as backward-looking in that the neutral band estimate is determined once a sequence of $d_{t\to k}$ is observed. Moreover, neutral band and arbitrage opportunities are modelled as occurring simultaneously. This approach is therefore omitting the timing of decisions made by market participants.

Consider the following sequence of events, which more closely resembles what goes on within a bank that aims to take advantage of arbitrage opportunities. Before entering any arbitrage activity, the market participant knows $b_{U,t}$ and $b_{L,t}$, possibly estimated by a risk-management department. That is, risk management ought to estimate $b_{U,t}$ and $b_{L,t}$ *before* any arbitrage transaction takes place at period t. The latter makes the last observed deviation from CIP the best starting point for producing a forecast, as opposed to the *theoretical* zero value. Moreover, note that the "counting" approach to estimation of w_t discussed above resembles the Value-at-Risk methodology.

Formally, let $\{\mathcal{F}_t\}_{t=1}^n$ be the filtration containing any information available up-to time t. Note that $\mathcal{F}_t$ contains both observable information in the data, as well as any unobservable that is inherent to the bank or the industry. Assume further that $\{d_{t\to k}\}_{t=1}^n$ is adapted to $\{\mathcal{F}_t\}_{t=1}^n$. Conditional on both the described timing framework and the "counting" approach, it is immediate that $b_{U,t}$ and $b_{L,t}$ satisfy

$$p\left(d_{t\to k} > b_{U,t\to k}|\mathcal{F}_{t-1}\right) = p\left(d_{t\to k} < b_{L,t\to k}|\mathcal{F}_{t-1}\right) = 0.025, \tag{2.6}$$

where $p\left(\cdot|\mathcal{F}_{t-1}\right)$ is a predictive distribution for $d_{t\to k}$, conditional on available information at $t-1$, $\mathcal{F}_{t-1}$. Note that this is equivalent to stating that $b_{U,t}$ and $b_{L,t}$ are particular values of a *density forecast*. Hence, a necessary condition to estimate w_t is having an estimate of the density forecast, $\widehat{p}\left(\cdot|\mathcal{F}_{t-1}\right)$. This is the task undertaken in the next section, after describing the data used in the paper.

2.2 Data

Data used in the empirical analysis are sampled on a weekly frequency, as in recent studies (Du et al., 2018). In particular, synchrony in collected data in measuring deviations from CIP is very important. If an investor were to compute said deviations with miss-matched data he would be making decisions described in Section 2.1 based on either a forward premium that is no longer available to take advantage of the interest rate differential, or vice-versa.[1]

To address the synchrony requirement, end-of-day data observed on Tuesday are used. Moreover, this reduces the number of observations that coincide with holidays, which take place mostly on Mondays or Fridays -which would limit, in turn, the use of daily data. It also avoids some biases observed in financial markets, such as "settlement Wednesdays" (Piazzesi,

[1] Taylor (1987) discusses the possible biases in measuring deviations from CIP at different times or dates.

2010).[2] Data are obtained from Bloomberg, unless otherwise stated.

Spot and 3-month forward exchange rate for the crosses GBP-USD, EUR-USD, and MXN-USD are used to compute the forward premium Φ_t. Interest rate differentials are computed using the interbank lending 3-month interest rates, which is the most common maturity used in the literature (Ivashina et al., 2015; Cenedese et al., 2017; Avdjiev et al., 2017; Sushko et al., 2016), Du et al. (2018).[3] This is the natural starting point as it embodies the costs that banks face to fund themselves. In particular, LIBOR in GBP, Euribor, and Mexican Equilibrium Interbank Rate (obtained from Banco de México) vis-á-vis LIBOR in USD are used.[4]

Sample covers from January 7th, 2003 to December 26th, 2017 for a total of 782 observations. In all cases, deviations from CIP are computed using (2.5). Since all data used throughout the paper have 3-month maturity (i.e. $k = 1/4$ hence $P_k = 100 \cdot 4$), in what follows the index "$\to k$" is omitted. The 3-month maturity is the one used in most of the previous studies that estimate w_t (Frenkel and Levich, 1975, 1977; Peel and Taylor, 2002). For ease of exposition, the following discussion is divided into three periods: pre-, during- and post-Global Financial Crisis (GFC).

US Dollar-Sterling Cross

Deviations from CIP in the GBP-USD FX market are shown in Figure 1 where it can be seen that, pre-GFC period, the CIP behaved according to the theory outlined above. During the GFC, which started in July 2007 and ended in 2010, the marked distinct behaviour of d is straight forwardly related to financial stress events such as: the dried liquidity conditions in Europe for mortgage-backed-assets in July 2007; the Bear Sterns buyout in March 2008; the Lehman Brothers failure in September 2008; and the AIG bail-out shortly after.

The Figure reflects how d returned only gradually to pre-GFC levels during 2009, thanks in large part to an unprecedented provision of liquidity by major central banks through different mechanisms. In the post-GFC period, beginning in 2010, the sovereign debt crisis in Europe is evident in the Figure. This event was a latent source of financial stress up to mid-2012. After this date, d is close to zero for a brief period between 2013 and 2015. Values for d_t observed near the date of the "Brexit" vote in June 2016 are the highest after the GFC. Note finally that, in said period, d is non-negative.

[2] Moreover, conducting test for "Day of the week effect" following Solnik and Bousquet (1990); Dubois and Louvet (1996); Berument and Kiymaz (2001); Gregoriou et al. (2004) suggests that only data collected in Tuesday is free from bias for all the crosses considered in this paper.

[3] By using weekly sampled data the percentage of lost data because is not registered in a given Tuesday is 1.8, 1.5, and 4.6 for the GBP, EUR and MXN, respectively. For daily data, in turn, the percentage of lost data is 3.1, 3.3, and 8.87.

[4] For analysis made with sovereign rates in the Appendix, 3-month maturity sovereign bonds in domestic currency are used. In particular, yields from the yield curve for the Gilts, German sovereign bond (obtained from both Bloomberg and the Bank of England), and Mexican CETES (obtained from Banco de México), vis-á-vis yields from the US Government yield curve.

US Dollar-Euro Cross

Deviations from CIP in the case of the EUR-USD shown in Figure 2 are markedly different from those of the GBP-USD during and after the GFC, regardless of both currencies having a "reserve currency" status. Indeed, in the pre-GFC period, d behaved according to the theory outlined above. Disruptions in the CIP are only noticeable after the summer of 2007. In fact, after this date, zero-valued deviations are rare. This is associated to the sovereign debt crisis in Europe.

The non-conventional monetary policy followed by the European Central Bank (ECB) in the 2010 - 2017 period is reflected in the behaviour of d. It is worth mentioning that, in addition to liquidity provision mechanisms, the ECB moved to imposing negative interest rates on its deposit accounts with commercial banks. This, in turn, may be causing an increasing trend starting in 2014 and peaking about the time of the "Brexit" vote. As in the GBP-USD case, most values of d after 2007 are non-negative.

US Dollar-Mexican Peso Cross

Unlike the GBP or the EUR, the MXN is not considered a "reserve currency". It is, however, the second most traded Emerging Market currency according to data from the BIS (2016) and it ranked first in the previous assessment in 2013. Figure 3 shows that, similarly to the GBP and the EUR, behaviour of d for the MXN was only disrupted in July 2007. The largest deviations from CIP, in absolute value, are recorded during the GFC, but the currency swap programme implemented by the Federal Reserve contributed to stabilising d during 2009. Deviations remained positive during and after 2010 with local peaks about the aforementioned Greek debt crisis.

The MXN has experienced several shocks during and post-GFC periods. Among these were the "Taper-tantrum" in mid-2013, the oil-price slump in the second half of 2014, and uncertainty around the prevailing trade conditions with the US and Canada, with the latter stemming from the US presidential election campaign. Noticeably, it has returned to theoretical levels since 2016 with a second local peak at the end of 2016. The return to levels predicted by theory is remarkable when compared to the GBP and the EUR cases where d seem to be diverging from zero.

3 Econometric Analysis

This section describes the econometric approach that is used to estimate w_t. First, a motivation for treating deviations from CIP within the neutral band as a Martingale process is provided. Indeed, the latter should serve as a benchmark for the validity of the estimate. Then, the dynamic models to estimate w_t are introduced. A discussion on the results from estimation and model evaluation concludes the section.

3.1 Dynamics within the Neutral Band

As discussed in Section 2.1 arbitrage is the underlying activity that corrects deviations from CIP. Data, however, rarely satisfy exactly (2.1), as shown in Figures 1-3. Thus, the focus here is to estimate b_U and b_L (i.e. w) defined in Section 2. To discipline dynamics of the deviations from CIP process, d, note that one-way arbitrage will be prevalent regardless of the level of the deviation, whereas round-trip arbitrage will only be triggered at period t if $d_t > b_{U,t|t-1}$ or $d_t < b_{L,t|t-1}$ from (2.6). Hence, as argued by Deardorff (1979), the dynamics of d within w are determined primarily by one-way arbitrage.

From asset pricing theory, in turn, if d is a Martingale process (i.e. the expected value of d_t conditional on information available in period s, with $s \leq t$, is d_s) then there exists a Martingale Measure and, hence, the market for d is arbitrage-free.[5] Should the latter be satisfied, data would not object the validity of the w estimate. Herein lies the main difference between this paper and the work from Peel and Taylor (2002). Instead of testing if $b_{U,t|t-1} < d_t < b_{L,t|t-1}$, for given (and fixed) values of $b_{U,t|t-1}$ and $b_{L,t|t-1}$, the focus here is on determining if the majority of observations are replicable by a model of a Martingale process.

3.2 Models

To determine if d may be replicated by a model of a Martingale process, the goodness of fit is assessed by estimating the posterior predictive distribution of the data. Then, w is given by b_U and b_L, which, in turn, are estimated as the $q = 0.025$ and $1 - q$ quantiles, respectively, of the one-step-ahead density forecast for d, $\widehat{p}\,(d_t|\mathcal{F}_{t-1})$. Moreover, b_U and b_L are time-varying, thus allowing to rationalise that restrictions on arbitrage are updated frequently in response to developments in financial markets, presumably at each period t.

Two popular models, used to parametrise a Martingale process, are the local level model (LLM) and the stochastic volatility model (SV) which, in turn, are discrete Euler approximations to processes represented by stochastic differential equations. The LLM has the attractive feature of relying on a very small number of parameters, while the SV is able to represent a large class of Martingale processes, as discussed in Shephard (2013). These models are means to obtain $\widehat{p}\,(d_t|\mathcal{F}_{t-1})$. The LLM is described in detail in West and Harrison (2006) and is stated as

$$y_t = m_t + \sigma_u u_t, \tag{3.1}$$

$$m_t = m_{t-1} + \sigma_\nu \nu_t, \tag{3.2}$$

where (3.1) is the observation equation, and (3.2) is the state equation, and where y_t is the first difference of d_t, $(u_t, \nu_t)'$ is a 2×1 vector of iid standard Gaussian random variables. The objects to be estimated are the parameter vector $\psi = (\sigma_u, \sigma_\nu)'$, and the latent process for the

[5] In particular, if the market is arbitrage-free, then there exists a Martingale measure with positive probabilities from Theorem 6.2 in Cvitanić and Zapatero (2004), page 193. By a contrapositive argument, if the process is not a Martingale, then there is no relevant Martingale measure, hence, the market is not arbitrage-free. The definition of Martingale Measure may be found in page 189 of the same reference.

time-varying mean m.

The SV, described in detail in Kim, Shephard, and Chib (1998), is given by

$$y_t = \exp(h_t/2)\,\varepsilon_t, \tag{3.3}$$

$$h_t = \mu + \phi\,(h_{t-1} - \mu) + \sigma_\eta \eta_t, \tag{3.4}$$

where (3.3) is the observation equation, (3.4) is the state equation, h is the (log) time-varying conditional variance, and $(\varepsilon_t, \eta_t)'$ is a 2×1 vector of iid standard Gaussian random variables. Within the SV the conditional variance process is modelled as a mean-reverting first-order autoregressive process with mean μ. Mean reversion of variance is an established fact in financial variables. Moreover, this process is assumed to be stationary, that is $0 < \phi < 1$. The objects to be estimated in the SV model are the 3×1 parameter vector $\theta = (\mu, \phi, \sigma_\eta)'$, and the latent process for the (log) time-varying conditional variance, h. Note in passing that the assumptions imply symmetry of the neutral band, but this may be relaxed.

The model used for estimating w in the literature is the threshold autoregressive model (TAR) from Tong (1990) and Granger and Teräsvirta (1993), hence, the natural candidate for a benchmark model. The TAR may be stated as,

$$d_t = \begin{cases} \alpha_m + \rho_m d_{t-1} + \sigma_\epsilon \epsilon_t & \text{if } \kappa_L < d_{t-1} < \kappa_U, \\ \alpha_u + \rho_u d_{t-1} + \sigma_\epsilon \epsilon_t & \text{if } d_{t-1} \geq \kappa_U, \\ \alpha_l + \rho_l d_{t-1} + \sigma_\epsilon \epsilon_t & \text{if } d_{t-1} \leq \kappa_L, \end{cases} \tag{3.5}$$

where κ_U and κ_L are the estimated upper- and lower-threshold levels, ϵ_t is a standard Normal random variable, and $(\alpha_j, \rho_j, \sigma_\epsilon)'$ is a parameter vector for the mid, upper, and lower regimes in (3.5). Note that the Martingale process arises as a particular case when $(\alpha_m, \rho_m, \sigma_\epsilon)' = (0, 1, \sigma_\epsilon)'$.

The density forecast function associated to LLM, SV, and TAR is Gaussian. Hence, estimates of the forecast mean and variance are needed. These objects are defined by

$$\widehat{p}\,(d_t|\mathcal{F}_{t-1}) = \mathcal{N}\left(d_{t|t-1}, V_{t|t-1}\right), \tag{3.6}$$

$$d_{t|t-1} = E\,(y_t|\mathcal{F}_{t-1}) + d_{t-1}, \tag{3.7}$$

$$V_{t|t-1} = Var\,(y_t|\mathcal{F}_{t-1})\,. \tag{3.8}$$

Since the TAR outcome depends on the estimates of $\kappa \in \{\kappa_U, \kappa_L\}$ and is estimated in levels, the moments of the density forecast function are indexed by $d_{t|t-1,\kappa}$ and $V_{t|t-1,\kappa}$.

3.3 Estimates

Estimation is made using an expanding window sample to account for the real-time feature of b_U and b_L. The training sample goes from January 7th 2003 to November 30th 2004, 100 observations in total. It should be noted that no events causing financial stress were registered through this period. Since inference is conducted by means of the density forecast defined in (3.6), estimates of ψ, m, θ and h are required. In turn, note that the posterior simulators detailed below produce ergodic sequences $\left\{\psi^{(j)}, m^{(j)}\right\}_{j=1}^{M}$, $\left\{\theta^{(j)}, h^{(j)}\right\}_{j=1}^{M}$ for each sample of

the expanding window. A discussion of each model, and the algorithms for estimation, are now provided.

LLM-MLE

The Maximum Likelihood estimates for the LLM (LLM-MLE) described in expressions (3.1)-(3.2) are a natural starting point for estimating a time-varying process as w. It is inexpensive, in terms of computation demands, and the recursive nature of the Kalman Filter allows to include information updates and forecast errors in subsequent forecasts. Since the outcome is a *point estimate* of the parameters and latent process, hence, trivially ergodic, results obtained do not entirely reflect the uncertainty around estimation.[6] The steps to obtain the w estimate and the likelihood are as follows:

1. Define the training sample $t = 1, \ldots, t_0 = 100$.
2. Obtain MLE point estimates of ψ and m using the R package *dlm* provided by Petris (2010).
3. Apply the Kalman filter using the *dlm* package, with ψ from previous step. Save the predicted value $d_{t|t-1}$ and its variance $V_{t|t-1}$ defined in (3.7) and (3.8).
4. Estimate $\widehat{p}\left(d_t|\mathcal{F}_{t-1}\right)$ by obtaining draws, $d^{(j)}_{t|t-1}$, from $\mathcal{N}\left(d_{t|t-1}, V_{t|t-1}\right)$, $j = 1, \ldots, M$, with $M = 10^4$.
5. Compute the log-likelihood of the next observation being contained in the forecast distribution as $LPL_t = 1/M \sum_{j=1}^{M} \ln \widehat{p}\left(d_t|d^{(j)}_{t|t-1}, V_{t|t-1}, \mathcal{F}_{t-1}\right)$.
6. Define $b_{U,t|t-1}, b_{L,t|t-1}$ as the 0.975 and 0.025 quantiles from (3.6), respectively, and $w_{t|t-1} = b_{U,t|t-1} - b_{L,t|t-1}$.
7. Add an observation to t_0 and go back to 2.

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

LLM-Bayesian

The Bayesian estimates for the LLM (LLM-Bayesian) described in expressions (3.1)-(3.2) are obtained through Gibbs-Sampling. As in the LLM-MLE case, this estimate possesses the advantages of being relatively parsimonious and those inherent to the Kalman Filter recursions. Unlike the MLE, however, this algorithm is able to account for the uncertainty around the estimates for ψ and m since the outcome is its posterior distribution. The cost

[6] Specification tests for the LLM-MLE models are constrained to the Ljung-Box Q statistic, since serial correlation is the only relevant violation of assumptions for inference. The latter, in view of the relatively large number of observations with respect to the number of parameters estimated. The test suggest that residuals associated with the MXN-USD cross are the only displaying serial correlation.

incurred comes in the form of computation demands. The (diffuse) prior distribution for the elements of ψ is Inverse-Gamma with hyperparameters 2 and 0.0001, denoted as $\psi \sim \mathcal{IG}(2, 0.0001)$.[7] It should be noted that the size of the sample guarantees that the likelihood function (or sampling distribution) will dominate the results, as opposed to estimations where prior distributions are highly relevant in view of a small sample. The steps to obtain the w estimate and the likelihood are as follows:

1. Define the training sample $t = 1, \ldots, t_0 = 100$.

2. Obtain the posterior distribution of ψ and m using the R package *dlm* provided by Petris (2010) and the Forward-Filter-Backward-Sampling (FFBS) from Carter and Kohn (1994) and detailed in Shumway and Stoffer (2017). Note that this is given by $\left\{\psi^{(j)}\right\}_{j=1}^{M}$ with $M = 10^4$ and a burn-in sample of 10^3. A relatively small burn-in sample may be enough to avoid the effects of initial conditions as long as distributions involved in estimation are not multi-modal, as suggested by Gelman et al. (2013).

3. Apply the Kalman filter using the *dlm* package, with each $\psi^{(j)}$ from previous step. Save the predicted value $d_{t|t-1}^{(j)}$ and its variance $V_{t|t-1}^{(j)}$ defined in (3.7) and (3.8).

4. Estimate $\widehat{p}\left(d_t|\mathcal{F}_{t-1}\right)$ by obtaining a draw from $\mathcal{N}\left(d_{t|t-1}^{(j)}, V_{t|t-1}^{(j)}\right)$ for each j.

5. Compute the log-likelihood of the next observation being contained in the forecast distribution as $LPL_t = 1/M \sum_{j=1}^{M} \ln \widehat{p}\left(d_t|d_{t|t-1}^{(j)}, V_{t|t-1}^{(j)}, \mathcal{F}_{t-1}\right)$.

6. Define $b_{U,t|t-1}, b_{L,t|t-1}$ as the 0.975 and 0.025 quantiles from (3.6), respectively, and $w_{t|t-1} = b_{U,t|t-1} - b_{L,t|t-1}$.

7. Add an observation to t_0 and go back to 2.

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

SV-FFBS

The SV-FFBS is the SV described in expressions (3.3)-(3.4) and is estimated through Markov Chain Monte Carlo (MCMC) methods proposed by Kim et al. (1998), using the same diffuse prior distributions and hyperparameters. In particular the prior distributions are given by $\sigma_\eta^2 \sim \mathcal{IG}(5, 0.05)$; $\phi \sim \mathcal{B}(20, 1.5)$ (the Beta distribution); and $\mu \sim \mathcal{N}(0, 10)$.[8] The prior distributions and hyperparameters are chosen to guarantee that $\sigma_\eta^2 > 0$ and $\phi \in (-1, 1)$.[9] Smoothed series for volatility estimated through FFBS is used to obtain forecast density.

[7] Graphical inspection of the posterior distribution of each element of ψ shows that chains mix relatively well, serial correlation is negligible beyond the 5th draw and distributions are uni-modal.

[8] Each element of θ shows that chains mix well. Moreover, serial correlation is negligible and posterior distributions are uni-modal.

[9] The parameter space for ϕ is bounded to guarantee that the volatility process is stationary, as suggested by theory. This is analogous to the assumptions on parameter estimates required for stationarity in GARCH(1,1) models in this regard.

An advantage of modelling volatility as a (stochastic) time-varying process is the model's capability of identifying periods of high uncertainty. A caveat of this method for producing forecasts is that every time a new observation is collected, the MCMC must be implemented to obtain the posterior distribution of h. The steps to obtain the w estimate and the likelihood are as follows:

1. Define the training sample $t = 1, \ldots, t_0 = 100$.

2. Obtain estimates for the posterior distributions of θ and h using the R package *stochvol* provided by Kastner (2016). Note that this is given by $\left\{\theta^{(j)}\right\}_{j=1}^{M}$ and $\left\{h_t^{(j)}\right\}_{j=1}^{M}$ with $M = 10^4$, a burn-in sample of 10^3, and $t = 1, \ldots, t_0$.

3. Estimate $\widehat{p}\left(d_t|\mathcal{F}_{t-1}\right)$ by obtaining a draw from $\mathcal{N}\left(d_{t|t-1}^{(j)}, V_{t|t-1}^{(j)}\right)$ for each j:

   (a) Compute $h_{t|t-1}^{(j)} = \mu^{(j)} + \phi^{(j)}\left(h_{t-1}^{(j)} - \mu^{(j)}\right) + \sigma_\eta^{(j)}\eta_t$, where η_t is a draw from $\mathcal{N}(0,1)$.

   (b) Compute $d_{t|t-1}^{(j)} = y_{t|t-1}^{(j)} + d_{t-1}$, where $y_{t|t-1}^{(j)} = \sqrt{V_{t|t-1}^{(j)}}\varepsilon_t$, $V_{t|t-1}^{(j)} = \exp\left(h_{t|t-1}^{(j)}\right)$ and ε_t is a draw from $\mathcal{N}(0,1)$, defined in (3.7) and (3.8).

4. Compute the log-likelihood of the next observation being contained in the forecast distribution as $LPL_t = 1/M \sum_{j=1}^{M} \ln \widehat{p}\left(d_t|d_{t|t-1}^{(j)}, V_{t|t-1}^{(j)}, \mathcal{F}_{t-1}\right)$.

5. Define $b_{U,t|t-1}, b_{L,t|t-1}$ as the 0.975 and 0.025 quantiles from (3.6), respectively, and $w_{t|t-1} = b_{U,t|t-1} - b_{L,t|t-1}$.

6. Add an observation to t_0 and go back to 2.

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

SV-APF

The SV-APF is also described in expressions (3.3)-(3.4) and is estimated through MCMC methods as in the case of SV-FFBS using the same prior distributions and hyperparameters. The posterior distribution of h is obtained using the Auxiliary Particle Filter (APF) proposed by Pitt and Shephard (1999). This methodology also possesses the advantages inherent to modelling volatility as a time-varying process. Unlike the SV-FFBS, SV-APF is able to more efficiently accommodate the arrival of new information since it only relies on the last estimate of the volatility. This is due to the forward recursions of the APF. The steps to obtain the w estimate and the likelihood are as follows:

1. Define the training sample $t = 1, \ldots, t_0 = 100$.

2. Obtain estimates for the posterior distribution of θ using the R package *stochvol* provided by Kastner (2016). Note that this is given by $\left\{\theta^{(j)}\right\}_{j=1}^{M}$ with $M = 10^4$, a burn-in sample of 10^3, and $t = 1, \ldots, t_0$. Compute the posterior mean for each parameter, obtaining $\bar{\theta}$.

3. Fixing parameters at $\bar{\theta}$, apply the APF and obtain $\left\{h_t^{(j)}\right\}_{j=1}^{M}$ with $M = 10^4$, and $t = 1, \ldots, t_0$.

4. Estimate $\widehat{p}\left(d_t|\mathcal{F}_{t-1}\right)$ by obtaining a draw from $\mathcal{N}\left(d_{t|t-1}^{(j)}, V_{t|t-1}^{(j)}\right)$ for each j:

   (a) Compute $h_{t|t-1}^{(j)} = \mu^{(j)} + \phi^{(j)}\left(h_{t-1}^{(j)} - \mu^{(j)}\right) + \sigma_\eta^{(j)}\eta_t$, where η_t is a draw from $\mathcal{N}(0,1)$.

   (b) Compute $d_{t|t-1}^{(j)} = y_{t|t-1}^{(j)} + d_{t-1}$, where $y_{t|t-1}^{(j)} = \sqrt{V_{t|t-1}^{(j)}}\varepsilon_t$, $V_{t|t-1}^{(j)} = \exp\left(h_{t|t-1}^{(j)}\right)$ and ε_t is a draw from $\mathcal{N}(0,1)$ defined in (3.7) and (3.8).

5. Compute the log-likelihood of the next observation being contained in the forecast distribution as $LPL_t = 1/M \sum_{j=1}^{M} \ln \widehat{p}\left(d_t|d_{t|t-1}^{(j)}, V_{t|t-1}^{(j)}, \mathcal{F}_{t-1}\right)$.

6. Define $b_{U,t|t-1}, b_{L,t|t-1}$ as the 0.975 and 0.025 quantiles from (3.6), respectively, and $w_{t|t-1} = b_{U,t|t-1} - b_{L,t|t-1}$.

7. Add an observation to t_0 and go back to 2.

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

TAR

The TAR is estimated by non-linear least squares (NLLS), as explained in Peel and Taylor (2002). As mentioned previously, this is the model that has been used to obtain an econometric estimate of w. Hence, it is a natural benchmark to assess the results from the set of models described above. Note that this model is estimated in levels with an autoregressive component of order 1 to mimic previous results in the literature.[10] The steps to obtain the w estimate and the likelihood are as follows:

1. Define the training sample $t = 1, \ldots, t_0 = 100$.

2. Obtain NLLS point estimates of $(\alpha_i, \rho_i, \sigma_\epsilon, \kappa)'$ using the R package *tsDyn* provided by Di Narzo et al. (2012). Here $i \in \{m, u, l\}$ as in (3.5).

3. Obtain the one-step ahead forecast $d_{t|t-1,\kappa}$ and its variance $V_{t|t-1,\kappa}$, defined in (3.7) and (3.8).

4. Estimate $\widehat{p}\left(d_t|\mathcal{F}_{t-1}\right)$ by obtaining draws, $d_{t|t-1,\kappa}^{(j)}$, from $\mathcal{N}\left(d_{t|t-1,\kappa}, V_{t|t-1,\kappa}\right)$, $j = 1, \ldots, M$, with $M = 10^4$.

5. Compute the log-likelihood of the next observation being contained in the forecast distribution as $LPL_{t_0} = 1/M \sum_{j=1}^{M} \ln \widehat{p}\left(d_t|d_{t|t-1,\kappa}^{(j)}, V_{t|t-1,\kappa}, \mathcal{F}_{t-1}\right)$.

[10] Regarding the specification tests, the Ljung-Box Q statistic for the whole sample suggest that only the GBP-USD cross fails to satisfy the no serial correlation assumption. This model-cross combination, however, is the only found in the literature.

6. Define $b_{U,t|t-1}, b_{L,t|t-1}$ as the 0.975 and 0.025 quantiles from (3.6), respectively, and $w_{t|t-1} = b_{U,t|t-1} - b_{L,t|t-1}$.

7. Add an observation to t_0 and go back to 2.

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

3.4 Results

General features of each type of model are immediately apparent in the w estimates across the board. In particular, estimates obtained from the LLM and TAR models are highly persistent. As discussed below for each cross, this is difficult to reconcile with both theory and financial stress events observed in the sample, either global or idiosyncratic. Estimates from SV models, in turn, are able to reflect the increase in transaction costs and risks associated with financial stress events. Moreover, these estimates point to arbitrage opportunities that are disperse through time, thus implying the presence of both one-way and round arbitrage. Clustering of the suggested arbitrage opportunities by the LLM and TAR models are implausible according to one-way arbitrage theory and reasonable assumptions on the works of the FX markets. Finally, persistence and width of the w estimates from the LLM and TAR are due to their inherent static variance estimates.[11]

US Dollar-Sterling Cross

Results show that w estimates for the GBP-USD cross obtained from the LLM and TAR are notably wide. These are displayed in Figure 4 left column, along with d.[12] Indeed, observed values of d_t at period t outside $w_{t|t-1}$ (marked with a vertical line) are clustered around the GFC (rows 1, 2 and 5). This is at odds with the surveyed literature, which has documented increases in uncertainty and transaction costs in the latter part of the sample. Moreover, the GBP-USD cross experienced episodes of considerable financial stress during the "Brexit" referendum in mid-2016. These models do not recognise the latter as a period where w increases.

Observed values of d_t outside $w_{t|t-1}$ obtained from the SV models, in turn, seem short-lived since they are more dispersed through time. The magnitude of the w estimate also reflects the financial stress originated in Europe in late 2010 and the "Brexit" eventsin June 2016, in addition to the GFC. The estimated time series for w is displayed in Figure 4 right column, and a summary is presented in Table 1. As in section 2.2, for ease of exposition the Table presents the range of values for w in periods pre-, during- and post-GFC.

[11] Results do not change if d is computed using sovereign bonds. These are included in the appendix.
[12] Figures 10-12 display the sub-sample 2014-2017.

Model/Period	Pre-GFC (Dec04-Jun07)	During-GFC (Jun07-Dec09)	Post-GFC (Jan10-Dec17)
LLM-MLE	10-12	10-57	37-53
LLM-Bayesian	11-14	11-58	38-54
SV-FFBS	5-22	8-356	2-52
SV-APF	6-16	9-558	2-69
TAR	9-11	9-54	35-51

Table 1: Estimated range for w in basis points in the GBP-USD FX market in each period.

In particular, results related to the pre-GFC period, from December 2004 to June 2007, point towards relatively similar highest estimates for w at around 12 basis points with the exception of the SV-FFBS. The lowest estimates are similar only within classes of models. LLM and TAR estimates are around 10 basis points, whereas SV's are around five basis points. During the GFC in the period from June 2007 to December 2009, the estimated lowest value of w is very similar across models, around nine basis points. The highest value, however, is again similar only within classes of models. LLM and TAR estimates are around 56 basis points, whereas SV's are above 350 basis points, hence they have a broader range.

Estimates for post-GFC period also show that, within class of models, results are similar. Interestingly, the LLM and TAR maximum estimates are very similar to those obtained for the period during the GFC in a range of 35-54 basis points. The latter would suggest that uncertainty and transaction costs are similar to those prevailing at the peak of the crisis. The SV models, in turn, estimate a range of 2-58 basis points, higher than pre-GFC, but nowhere near the peaks attained then.

Previous estimates of the neutral band for the GBP-USD cross were described in the literature review above. Results obtained for the pre-GFC period are similar. In particular the SV models, which estimate a neutral band width of a minimum of 5 and a maximum of 22 basis points, are similar to Taylor (1987) and Clinton (1988) and smaller than Frenkel and Levich (1975, 1977). The estimated w in these papers and in the present work are well below early estimates from Keynes and Enzig, in line with the development of the relevant markets and communication technology.

US Dollar-Euro Cross

The w estimates for the EUR-USD are displayed in Figure 5 left column along with the measured d. As in the GBP-USD estimates, the LLM and TAR yield considerably wider estimates than those obtained from the SV. Notably, the sovereign debt crisis experienced in Europe through 2010 and peaking in the last quarter of 2012, is recognised by all models. The LLM and TAR, however, only suggest arbitrage opportunities during episodes of considerable financial stress.

As opposed to results from the SV, there is clustering of the measured d outside of w by the LLM and TAR. This suggests that arbitrage opportunities are long-lived, a feature that is difficult to reconcile with reality. The estimated magnitudes of w are shown in Figure 5 right column. The LLM and TAR estimates are rather persistent, unlike those obtained from the SV. This, in turn, reveals that, as in the GBP-USD analysis, the LLM and TAR do not account fully for changes in transaction costs or uncertainty. The w estimates yielded by the SV seem to track closely episodes of heightened stress in financial markets. The latter is confirmed by the range of estimates presented in Table 2.

Results suggest that the highest estimates from the LLM and TAR during- and post-GFC periods are remarkably similar, at around 60 basis points. But this suggests that, between 2010 and 2017, the FX market for EUR-USD was subject to financial stress similar to that in the peak of the GFC. The latter is at odds with findings in the surveyed literature and the extraordinary low levels of interest rates across advanced economies.

In contrast, the highest w estimates yielded by the SV are reduced from 363 to 104 basis points, still reflecting a high degree of financial stress related to the sovereign debt crisis but considerably lower than that experienced during the GFC. Unfortunately, there are no estimates of the neutral band for the EUR-USD cross, hence it is not possible to directly compare this results.

Model/Period	Pre-GFC (Dec04-Jun07)	During-GFC (Jun07-Dec09)	Post-GFC (Jan10-Dec17)
LLM-MLE	7-8	7-60	42-56
LLM-Bayesian	10-11	10-62	42-58
SV-FFBS	5-14	7-363	5-121
SV-APF	7-12	8-253	5-107
TAR	8-9	8-56	40-53

Table 2: Estimated range for w in basis points in the EUR-USD FX market in each period.

US Dollar-Mexican Peso Cross

In Figure 6, left column, it is apparent that observations of d outside w for the MXN-USD cross reflect two global financial shocks, during 2007-2009 and the last quarter of 2010 periods. Moreover, all models suggest arbitrage opportunities around the end of November 2016, possibly associated with the presidential election campaign in the US.

In addition to these episodes, the SV models identify arbitrage opportunities during the Brexit referendum in mid-2016, and the "tapper tantrum" in mid-2013. As in the GBP-USD and EUR-USD estimations, the LLM and TAR estimate persistently large w with clustered apparent arbitrage opportunities, whereas the SV are able to accommodate bouts of uncertainty and suggest disperse arbitrage opportunities. Furthermore, the SV models replicate both the financial stress associated with the MXN real depreciation after the slump in oil

prices in 2014, and uncertainty around the outcome of the revamped trade agreement with the US and Canada in mid-2017.

Figure 6, right column, displays the magnitudes of w and Table 3 the range of values for the three sub-periods as in the GBP- and EUR-USD crosses. The SV maximum estimates of w during the GFC are at least 300 basis points larger than those for the 2010-2017 period, from 846 to 324 basis points. Ranges for the LLM and TAR are remarkably similar in the post-GFC periods, as in previous cases. Unfortunately, as in the case of the EUR-USD cross, there are no previous estimates of the neutral band for comparison.

Model/Period	Pre-GFC (Dec04-Jun07)	During-GFC (Jun07-Dec09)	Post-GFC (Jan10-Dec17)
LLM-MLE	71-96	68-148	108-141
LLM-Bayesian	71-97	68-151	109-143
SV-FFBS	17-125	22-846	30-331
SV-APF	28-96	28-711	36-370
TAR	63-83	60-145	103-137

Table 3: Estimated range for w in basis points in the MXN-USD FX market in each period.

3.5 Model Evaluation

It was shown that, in general, the time-varying-mean/constant-variance models suggest a neutral band that is seemingly wide throughout the sample. The upside in terms of risk-management of these estimates is that they may provide more a conservative w. These entail a large opportunity cost in terms of missed arbitrage opportunities, however. Moreover, the latter is obtained at the cost of not recognising financial stress periods, which should be reflected in increases of w.[13] The time-varying variance models, in contrast, are able to account for changes in uncertainty and provide a narrower and less persistent w. As discussed above, this is consistent with short-lived arbitrage opportunities.

Cumulative Log-Predictive Likelihood

To strike a balance between having a conservative w, without sacrificing precision in terms of identifying financial stress periods, an in-sample evaluation and a density forecast evaluation are conducted using the methods suggested by Geweke and Amisano (2010). Both evaluations rely on computing the *cumulative* log-predictive-likelihood. Recall LPL_t from Section 3.3, and define

$$CLPL_t = \sum_{j=t_0}^{t} LPL_j, \ t = t_0, \ldots, n.$$

[13] In line with the poor predictive properties of TAR models documented, for example, in Dacco and Satchell (1999).

Results from the in-sample evaluation are contained in Table 4, which displays $CLPL_n$ for each currency-model combination.[14] Immediate notable features are, first, that in all crosses the best model is the SV-APF. Second, the next best model is the LLM-Bayesian for the GBP-USD and EUR-USD crosses, while the SV-FFBS is the second best model for the MXN-USD cross. Third, the TAR used previously in the literature is considerably inferior in terms of fitting the data. Finally, results suggest that accounting for the uncertainty around the estimated parameters and latent process plays an important role in fitting the data, as reflected by the fact that models estimated by MCMC methods dominate.

	GBP-USD	EUR-USD	MXN-USD
LLM-MLE	56.70556	-256.0021	-491.635
LLM-Bayesian	306.6930+	251.1586+	-418.3470
SV-FFBS	223.1345	19.93174	-271.9914+
SV-APF	529.8467*	413.4730*	-174.0033*
TAR	-2810.025	-4409.585	-3766.946

Table 4: $CLPL_n$. * Refers to the combination with highest likelihood. + Is the second highest.

Log-Score

Density forecast evaluation is conducted through a log-score, LS_t, statistic. As detailed in Geweke and Amisano (2010), density forecast comparison is feasible regardless of the inference procedure as long as a time series for $CLPL_t$ is recovered. Since the comparison is pair-wise, a two-step evaluation is made for each FX market following these authors. Define

$$LS_t^{\mathcal{M}} = CLPL_t^{\mathcal{M}} - CLPL_t^{TAR}. \tag{3.9}$$

$$LS_t^{*} = CLPL_t^{(1)} - CLPL_t^{(2)}. \tag{3.10}$$

where $\mathcal{M} \in \{\text{LLM-MLE, LLM-Bayesian, SV-FFBS, SV-APF}\}$ and $CLPL_t^{(i)}$ is the cumulative predictive log-score corresponding to the model with the *i-th* highest $CLPL_n$. In the first step, a comparison is made based on $CLPL_n$, and $\left\{LS_t^{\mathcal{M}}\right\}_{t=t_0}^{n}$ computed in (3.9). For the second step, the two models with highest $CLPL_n$ are compared using $\{LS_t^{*}\}_{t=t_0}^{n}$ computed with (3.10).

Results for the GBP-USD cross as displayed in rows 1 and 2 in Figure 7 correspond to the first step. These suggest that both Martingale parametrisations out-perform the TAR model in terms density forecasts throughout the sample, since $LS_t^{\mathcal{M}}$ is positive for all t and has a positive slope. In particular, the spike in $\left\{LS_t^{\mathcal{M}}\right\}_{t=t_0}^{n}$ around the GFC, suggest that both the LLM and SV are more able to account for financial stress in the GBP-USD cross. The last row in Figure 7 displays the comparison between the LLM-Bayesian and SV-APF. The latter yields better density forecasts, hence, better estimates for w in the sample, except for one

[14] Table 7 in the Appendix displays results for the sovereign rates case.

observation in mid-2005, one in mid-2007 and one at the last quarter of 2008. The latter is possibly related to GFC, the Lehman Brothers' failure in particular.[15]

As in the GBP-USD cross, results for the EUR-USD FX cross all Martingale parametrisations and outperform the TAR, as shown in rows 1 and 2 in Figure 8. Even though the TAR model is outperformed consistently, as implied by the positive slope of the plots, it is interesting to note that two dates seem to explain the overwhelming result. First, the aforementioned Lehman failure during the GFC and, second, the peak of the sovereign debt crisis in Europe in 2012. The last row in Figure 8 displays the comparison between the LLM-Bayesian and the SV-APF. In particular, it shows that w is best estimated by the SV-APF for the sample 2009-2017. That is, for any financial stress event distinct from the GFC, the SV-APF is unambiguously the best candidate to estimate w. For the sample 2005-2007 the LLM-Bayes does a better job, possibly due to the fact that swings in d were more prevalent in this period (see Figure 5). Also noteworthy, the LLM-Bayesian has a better fit for w in dates that displayed a more acute financial stress such as the summer of 2007 and late 2008, both related to the GFC.[16]

Figure 9 shows that, as in previous crosses, all Martingale parametrisations produce better density forecasts, and, hence, better estimates of w than the TAR for the MXN-USD cross. Interestingly, it is only after the GFC that the TAR is clearly outperformed by the LLM and SV. As in the EUR-USD case, most of the difference of these models with respect the TAR is explained by the Lehman failure in the last quarter of 2008, during the GFC. Interestingly, instead of specific dates, the whole period related to the sovereign debt crisis stemming from Europe in 2011-2012 seems to explain a large portion of the under-performance of the TAR. This may be due to the large estimates for w during that period. The last row in Figure 9 displays the comparison between SV-APF and SV-FFBS. The latter yields better estimates of w across the sample, but the difference in performance is notably small. Among the main episodes explaining said difference is an apparent arbitrage opportunity arising in late 2005 (see Figure 6), when the interest rates in the Mexican interbank markets began a steep decline responding to monetary policy easing.[17]

Probability Integral Transformation

These results are subject to the shortcomings of using the $LS_t^{\mathcal{M}}$ statistic to evaluate forecast performance. In particular, as noted by Geweke and Amisano (2010), said statistic is a local measure of performance and is only a relative measure, hence, conditional on choosing a benchmark model. To address these shortcomings, results for the probability integral transformation (PIT) are included. Model assessment based on the PIT is not local. Further

[15] As shown in the Appendix Figure 22, the w estimate for the sovereign d is, with no exception, best fitted by the SV-APF.

[16] The SV-FFBS has a better fit for the sovereign w, as shown in Figure 23 in the appendix.

[17] Results vary considerably during the GFC in the sovereign w yielding the LLM-Bayesian as the best model for the period, as shown in Figure 24 in the appendix.

computations are required, since this measure is based on the (empirical) cumulative distribution function (CDF) of the forecasts (see discussion about expression (15) in the reference of Geweke and Amisano for further details). In particular, let $\mathbb{I}(\cdot)$ be an indicator function, then the (empirical) CDF is given by

$$\widehat{F}(d_t) = 1/M \sum_{j=1}^{M} \mathbb{I}\left(d_{t|t-1}^{(j)} \in (-\infty, d_t)\right).$$

Moreover, let $\Phi(x)$ be $\mathcal{N}(0,1)$ computed at x. Note that a necessary condition for the test statistic to work is ergodicity of the sequences $\left\{\psi^{(j)}, m^{(j)}\right\}_{j=1}^{M}$ and $\left\{\theta^{(j)}, h^{(j)}\right\}_{j=1}^{M}$, otherwise the CDF should be conditional on the uncertainty inherent to the estimated parameters and latent processes. The model assessment is then based on testing whether $\widehat{f}(d_t) = \Phi^{-1}\left(\widehat{F}(d_t)\right)$ is *iid* $\mathcal{N}(0,1)$.

Table 5 displays the *p-value* for three commonly used normality tests: Chi-squared goodness of fit test, Kolmogorov-Smirnov test, and Shapiro-Wilk test. It is worth noting that, despite the (very) low *p-values*, these should be indicative of how far the empirical distribution $\widehat{F}(d_t)$ is from the distribution under the null hypothesis. For ease of comparison, the highest *p-value* for each cross and each test is marked in the Table.

The PIT confirms the superior forecast performance of the SV for all crosses and all tests. In particular, the PIT suggests that the SV-FFBS is superior to the SV-AFP, with the exception of the GBP-USD tested with the Kolmogorov-Smirnov test. This result is in line with the described nature of both the $LS_t^{\mathcal{M}}$ statistic and the PIT. In particular, the SV-APF is more efficient in updating the information set to produce a forecast, hence, it uses new information in a "local" sense. The SV-FFBS, however, uses all information contained in the sample each time a new observation arrives, hence, it makes use of information in a "global" sense. Overall, the results from model evaluation suggest that SV-APF is best-suited to obtain forecasts allocating relatively more weight to recent information. The SV-FFBS, however, is best-suited to include all information contained in the sample.

4 Concluding Remarks

This paper undertakes the estimation of the neutral band around deviations from CIP. Seemingly large deviations have been observed in the post-Global Financial Crisis period in the GBP-USD, EUR-USD and MXN-USD FX markets. To assess whether these deviations are, in fact, arbitrage opportunities, as theory would suggest, an estimate of the neutral band is needed. The approach differs from previous work in that, if arbitrage opportunities are indeed absent, then a parametrisation of a Martingale process should be able to replicate deviations from CIP.

Results from model evaluation suggest that, for all crosses and cases, the stochastic volatility models, a particular Martingale process parametrisation, yield superior neutral band estimates. Several features suggest that these estimates have better properties than the alterna-

tives. First, they have a superior performance in- and out-of sample. Second, they are able to account for a wider neutral band in periods of financial stress. Third, the neutral band is wider during the Global Financial Crisis than in the rest of the sample, whereas competing models estimate similar values for the 2007-2017 period. Finally, the neutral band estimate implies arbitrage opportunities scattered through time, hence, short-lived. In contrast, the rival models imply clustered, and, hence, long-lived arbitrage opportunities.

Some caveats in interpreting the results are in order. First, estimates of the neutral band are based on the one-step-ahead density forecast of a latent process. As such, it is not possible to evaluate these estimates vis-á-vis observed data. It may be necessary to expand the number of econometric models and techniques to assess if the approach proposed here to estimate the neutral band may be improved upon. Second, results for the MXN may require refinement in that, unlike the GBP and the EUR, it is not considered a "reserve" currency.

Modelling distinct FX markets may require further experiments with econometric tools, such as a time-varying multivariate model, as the literature has aimed to model deviations from CIP with a single-variable model. Future research should aim to consider a number of alternative empirical models and data sets, which, in turn, may allow to condition on the idiosyncratic features of each currency.

	GBP-USD	EUR-USD	MXN-USD
	Chi-squared goodness of fit		
LLM-MLE	0	0	0
LLM-Bayesian	0	0	0
SV-FFBS	$5.44 \times 10^{-172}*$	$8.96 \times 10^{-139}*$	$1.45 \times 10^{-45}*$
SV-APF	0	0	6.35×10^{-92}
TAR	0	0	0
	Kolmogorov-Smirnov		
LLM-MLE	0	0	4.22×10^{-15}
LLM-Bayesian	0	0	0
SV-FFBS	0.111*	0.413*	0.884*
SV-APF	0.024	0.312	0.561
TAR	0	0	1.11×10^{-16}
	Shapiro-Wilk		
LLM-MLE	1.34×10^{-33}	5.06×10^{-31}	1.53×10^{-28}
LLM-Bayesian	0	0	0
SV-FFBS	$4.26 \times 10^{-5}*$	$5.35 \times 10^{-3}*$	0.453*
SV-APF	2.65×10^{-9}	2.03×10^{-8}	0.204
TAR	1.02×10^{-31}	4.38×10^{-28}	5.84×10^{-26}

Table 5: PIT. * Refers to the model with highest *p-value* for each cross and each test. Null-Hypothesis is normality.

References

Avdjiev, S., W. Du, C. Koch, and H. S. Shin (2017): "The dollar, bank leverage and the deviation from covered interest parity," *BIS Working Papers*, 592.

Baba, N. and F. Packer (2009): "Interpreting deviations from covered interest parity during the financial market turmoil of 200708," *Journal of Banking & Finance*, 33, 1953 – 1962, financial Globalisation, Risk Analysis and Risk Management.

Berument, H. and H. Kiymaz (2001): "The day of the week effect on stock market volatility," *Journal of Economics and Finance*, 25, 181–193.

BIS (2016): *Triennial Central Bank Survey: Foreign exchange turnover in April 2016*.

Borio, C. E., R. N. McCauley, P. McGuire, and V. Sushko (2016): "Covered interest parity lost: understanding the cross-currency basis," *BIS Quarterly Review*, September.

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

Bush, G. (2019): "Bank foreign currency funding and currency markets: the case of Mexico post GFC," *Banco de México Working Papers*, 2019-01.

Carstens, A. G. (1985): *A Study of the Mexican Peso Forward Exhcange Rate Market. PhD Dissertation*, University of Chicago.

Carter, C. K. and R. Kohn (1994): "On Gibbs sampling for state space models," *Biometrika*, 81, 541–553.

Cenedese, G., P. Della-Corte, and T. Wang (2017): "Currency Mispricing and Dealer Balance Sheets," *Mimeo*.

Claessens, S. and M. A. Kose (2018): "Frontiers of macrofinancial linkages," *BIS Working Papers*.

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

Cvitanić, J. and F. Zapatero (2004): *Introduction to the Economics and Mathematics of Financial Markets*, MIT press.

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, 351–364.

Di Narzo, A., J. Aznarte, and M. Stigler (2012): "Nonlinear time series models with regime switching," *R package version (http://cran. us. rproject. org/web/packages/tsDyn/)*.

Du, W., A. Tepper, and A. Verdelhan (2018): "Deviations from Covered Interest Rate Parity," *The Journal of Finance*, Forthcoming.

Dubois, M. and P. Louvet (1996): "The day-of-the-week effect: The international evidence," *Journal of Finance an Banking*, 20, 1463–1483.

Duffie, D. (2010): *Dynamic asset pricing theory*, Princeton University Press.

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

Frenkel, J. A. (1973): "Elasticities and the Interest Parity Theory," *Journal of Political Economy*, 81, 741–747.

Frenkel, J. A. and R. M. Levich (1975): "Covered Interest Arbitrage: Unexploited Profits?." *Journal of Political Economy*, 83, 325 – 338.

——— (1977): "Transaction Costs and Interest Arbitrage: Tranquil versus Turbulent Periods." *Journal of Political Economy*, 85, 1209 – 1226.

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin (2013): *Bayesian Data Analysis*, Chapman and Hall/CRC.

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

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

Gregoriou, A., A. Kontonikas, and N. Tsitsianis (2004): "Does the Day of the Week Effect Exist Once Transaction Costs Have Been Accounted for? Evidence From the UK," *Applied Financial Economics*, 14, 215–220.

Hernandez, J. R. (2014): "Peso-Dollar Forward Market Analysis: Explaining Arbitrage Opportunities during the Financial Crisis," *Banco de México Working Papers*, 2014-09.

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

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

Kastner, G. (2016): "Dealing with Stochastic Volatility in Time Series Using the R Package stochvol," *Journal of Statistical Software*, 69, 1–30.

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

Khor, H. E. and L. Rojas-Suarez (1991): "Interest Rates in Mexico: The Role of Exchange Rate Expectations and International Creditworthiness," *No 91/12 IMF Staff Papers*, 850–871.

Kim, S., N. Shephard, and S. Chib (1998): "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *The Review of Economic Studies*, 65, 361–393.

Levich, R. M. (2012): "FX counterparty risk and trading activity in currency forward and futures markets," *Review of Financial Economics*, 21, 102–110.

——— (2017): "CIP: Then and Now, A Brief Survey of Measuring and Exploiting Deviations from Covered Interest Parity." *Remarks prepared for the BIS Symposium: CIP - RIP?*

Liao, G. (2016): "Credit migration and covered interest rate parity," *Project on Behavioral Finance and Financial Stability Working Paper Series (2016-07)*.

Mancini-Griffoli, T. and A. Ranaldo (2011): "Limits to arbitrage during the crisis: funding liquidity constraints and covered interest parity," *Swiss National Bank Working Paper*.

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, 51 – 75.

Petris, G. (2010): "Dealing with Stochastic Volatility in Time Series Using the R Package stochvol," *Journal of Statistical Software*, 36, 1–16.

Piazzesi, M. (2010): "Affine Term Structure Models," in *Handbook of Financial Econometrics: Tools and Techniques*, ed. by Y. Aït-Sahalia and L. P. Hansen, North-Holland, vol. 1, 691 – 766.

Pitt, M. K. and N. Shephard (1999): "Filtering via Simulation: Auxiliary Particle Filters," *Journal of the American Statistical Association*, 94, 590–599.

Shephard, N. (2013): "Martingale unobserved component models," *Manuscript, University of Oxford*.

Shumway, R. H. and D. S. Stoffer (2017): *Time Series Analysis and Its Applications*, Springer.

Solnik, B. and L. Bousquet (1990): "Day-of-the-week effect on the Paris Bourse," *Journal of Finance an Banking*, 14, 461–468.

Sushko, V., C. E. Borio, R. N. McCauley, and P. McGuire (2016): "The failure of covered interest parity: FX hedging demand and costly balance sheets," *BIS Working Papers*, 590.

Taylor, M. P. (1987): "Covered Interest Parity: A High-Frequency, High-Quality Data Study," *Economica*, 54, 429–438.

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

West, M. and J. Harrison (2006): *Bayesian forecasting and dynamic models*, Springer Science & Business Media.

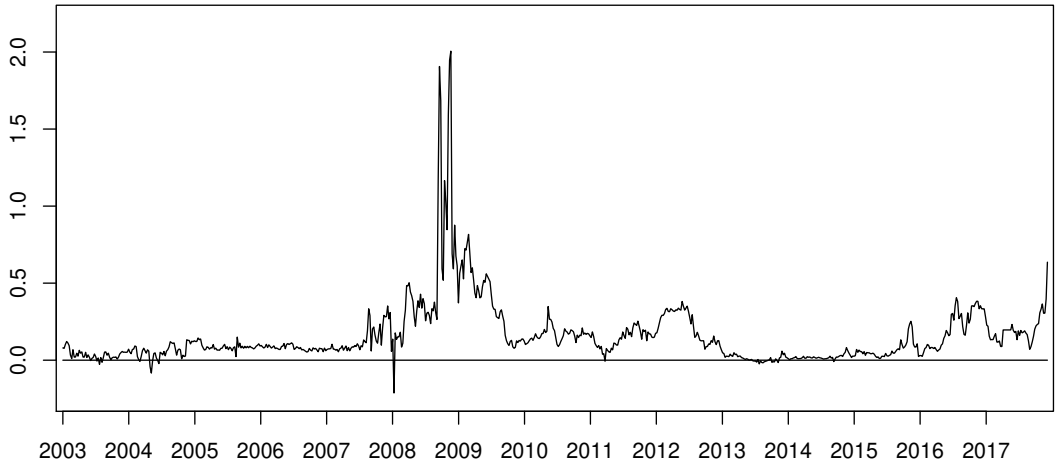


Figure 1: Deviations from CIP in percentage points GBP-USD. Computed with 3-month interbank rates. Sample: January 7th 2003 to December 26th 2017. Own calculations for weekly data. Source: Bank of England and Bloomberg.

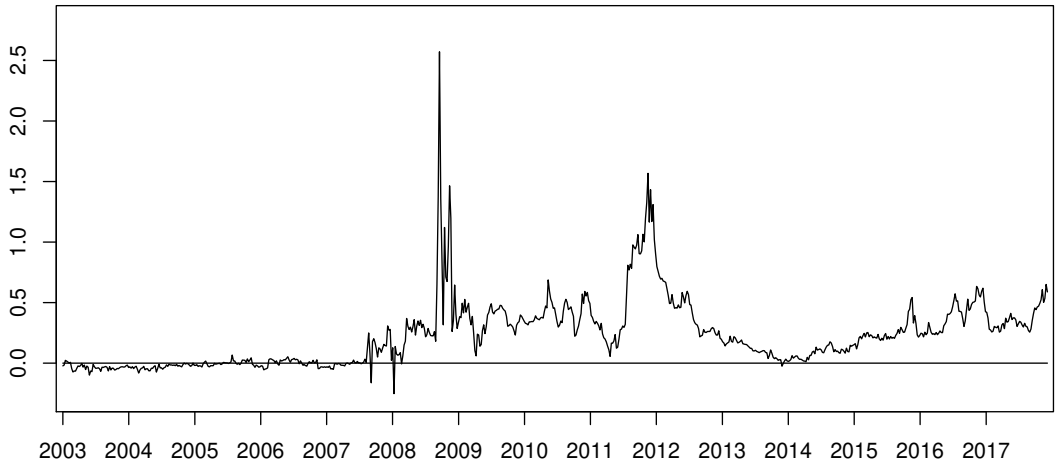


Figure 2: Deviations from CIP in percentage points EUR-USD. Computed with 3-month interbank rates. Sample: January 7th 2003 to December 26th 2017. Own calculations for weekly data. Source: Bloomberg.

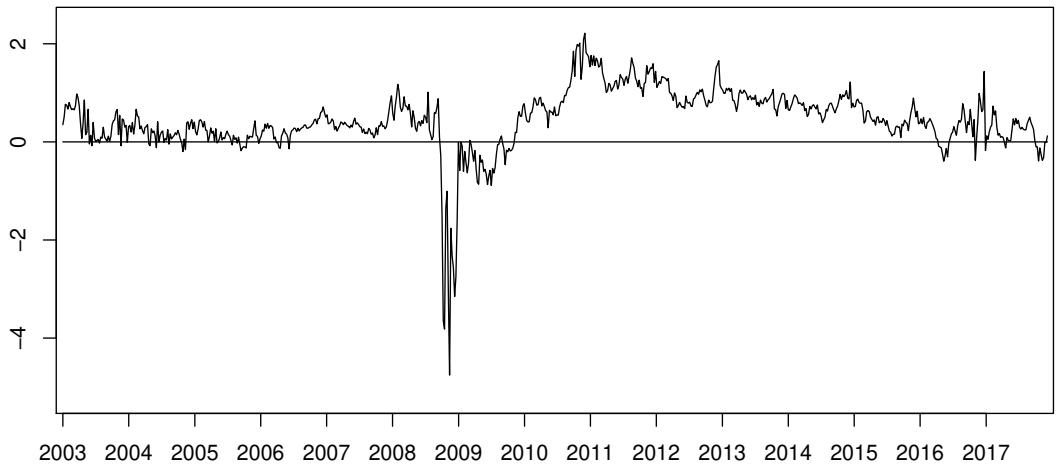


Figure 3: Deviations from CIP in percentage points MXN-USD. Computed with 3-month interbank rates. Sample: January 7th 2003 to December 26th 2017. Own calculations for weekly data. Source: Banco de México and Bloomberg.

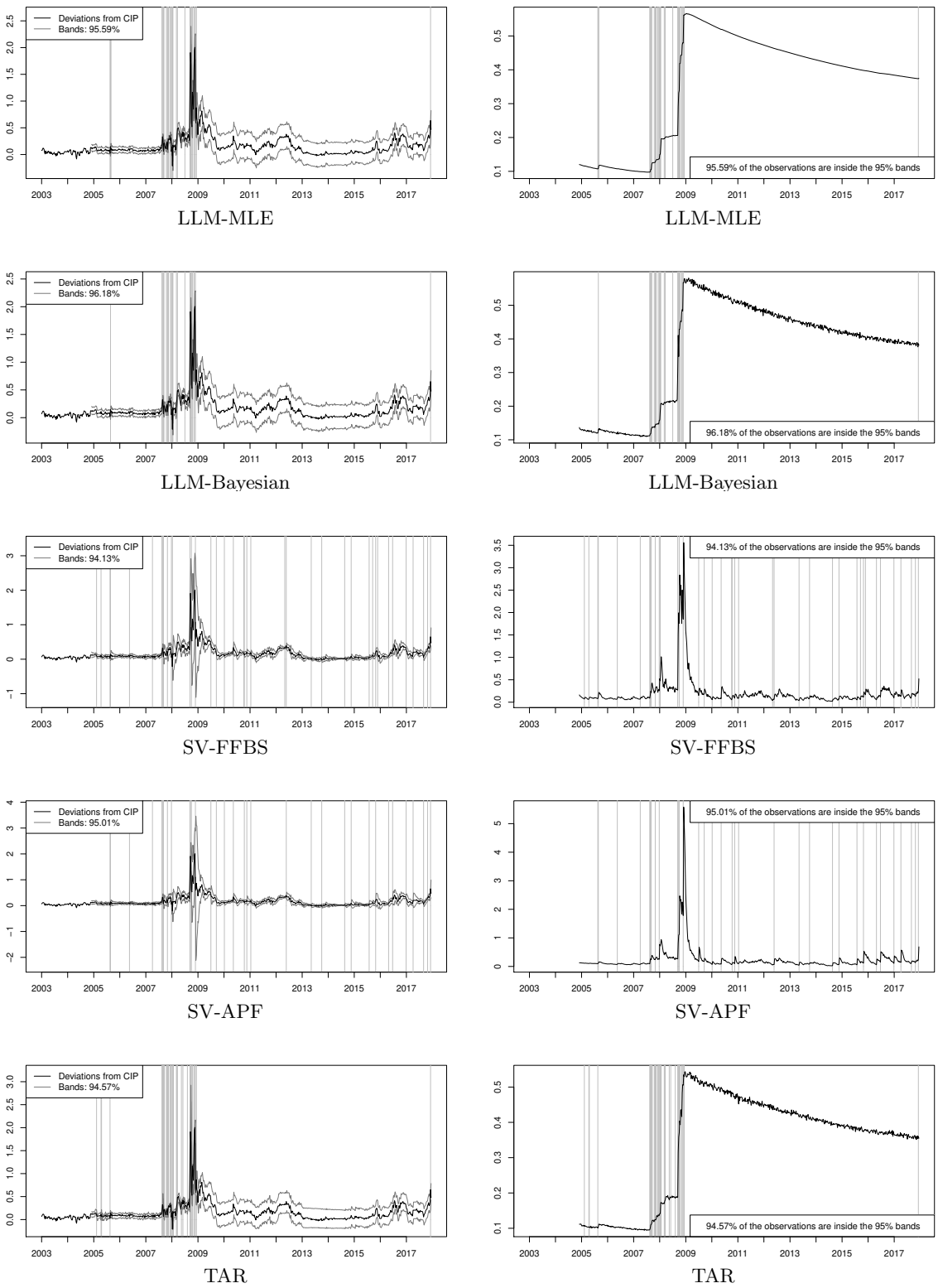


Figure 4: Deviations from CIP in percentage points and neutral band estimates GBP-USD (left column)and estimated w in percentage points (right column). Vertical lines are observations outside w. Label contains the percentage of observations within w.

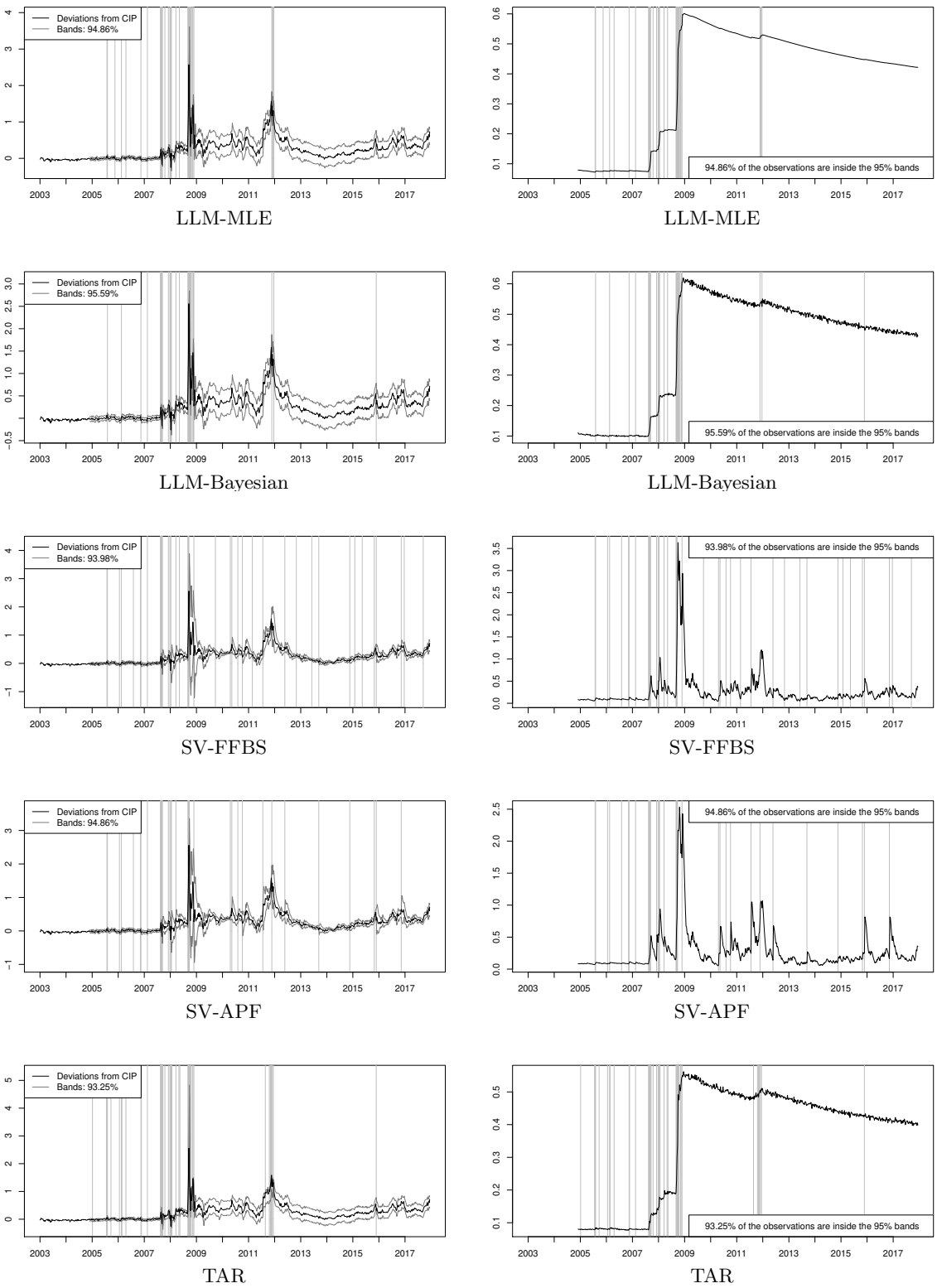


Figure 5: Deviations from CIP in percentage points and neutral band estimates EUR-USD (left column)and estimated w in percentage points (right column). Vertical lines are observations outside w. Label contains the percentage of observations within w.

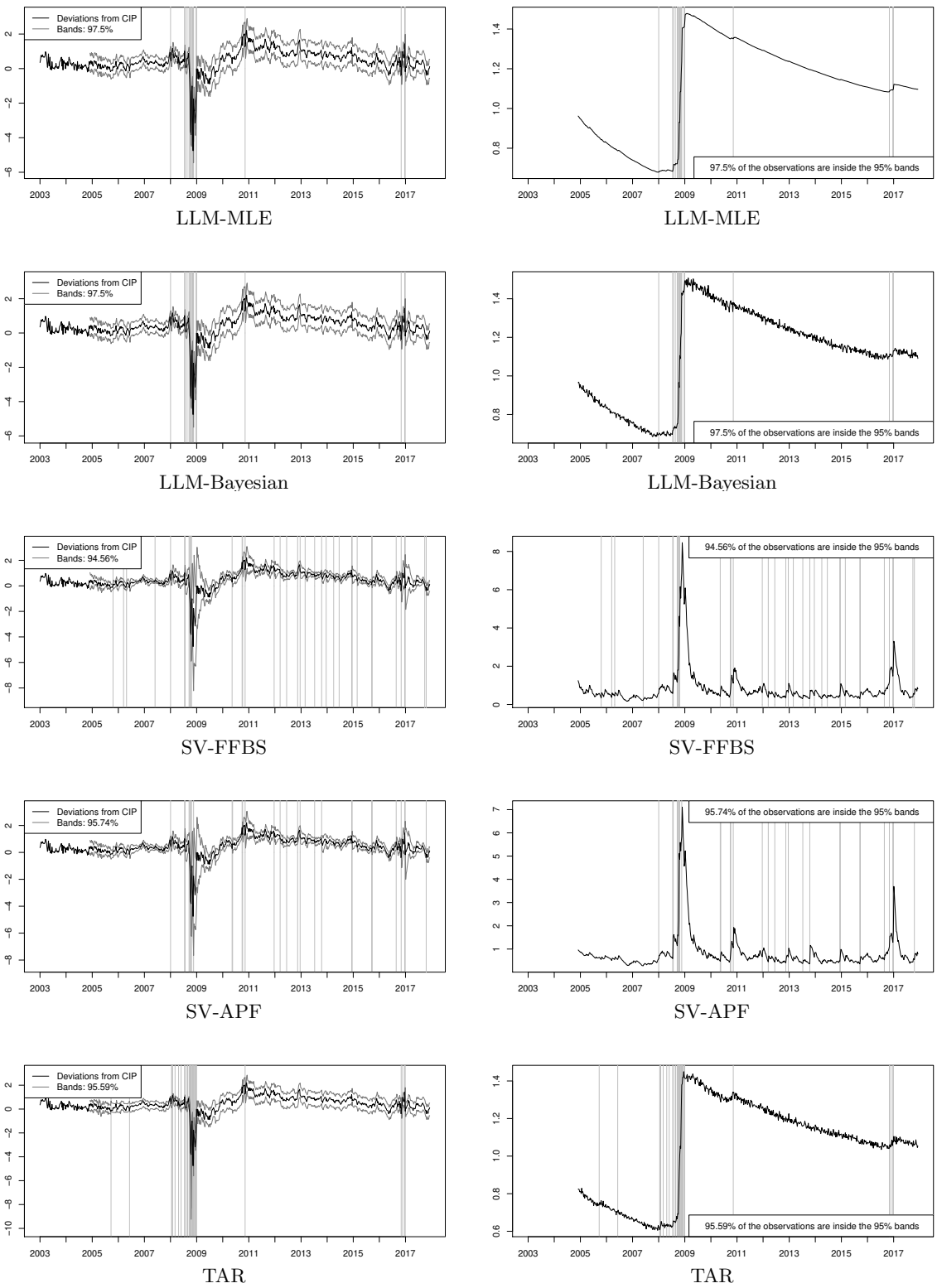
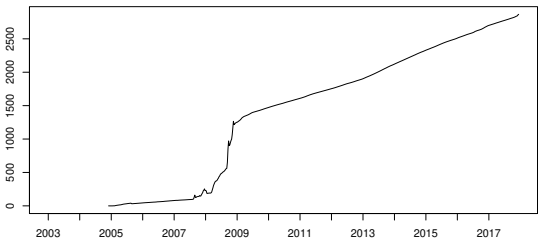
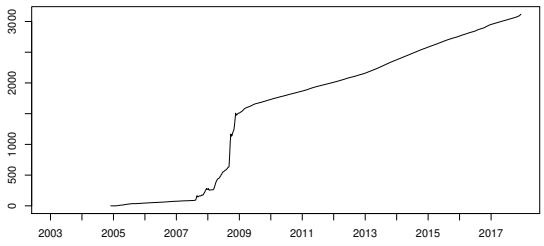


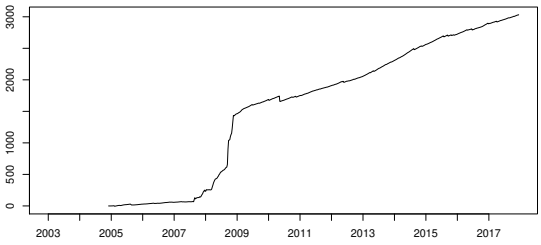
Figure 6: Deviations from CIP in percentage points and neutral band estimates MXN-USD (left column)and estimated w in percentage points (right column). Vertical lines are observations outside w. Label contains the percentage of observations within w.



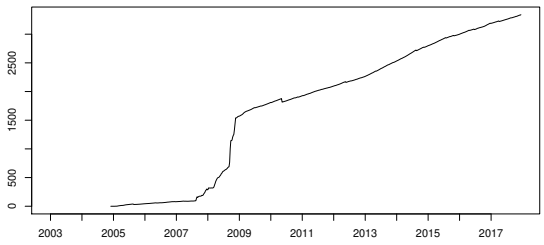
LLM-MLE



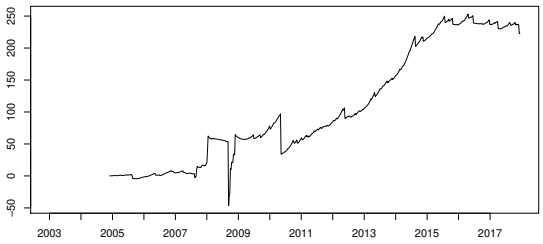
LLM-Bayesian



SV-FFBS

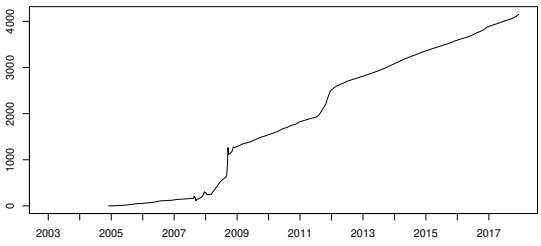


SV-APF

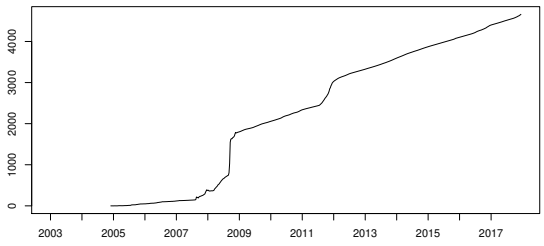


SV-APF/LLM-Bayesian

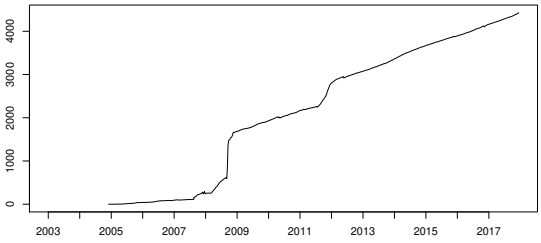
Figure 7: $LS_t^{\mathcal{M}}$ and LS_t^*, GBP-USD.



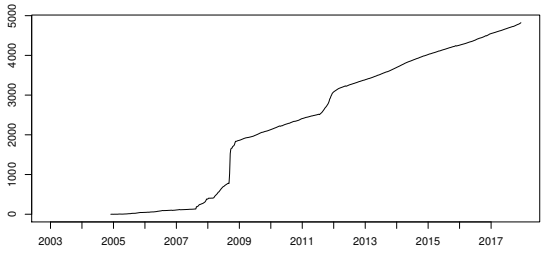
LLM-MLE



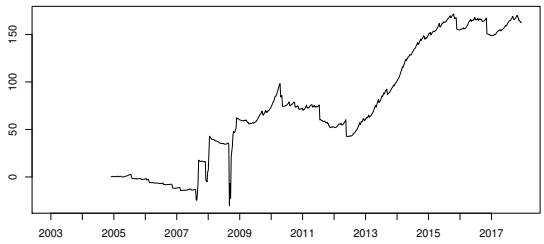
LLM-Bayesian



SV-FFBS

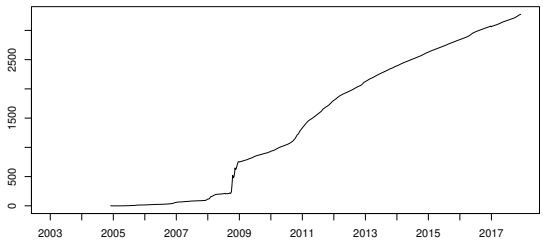


SV-APF

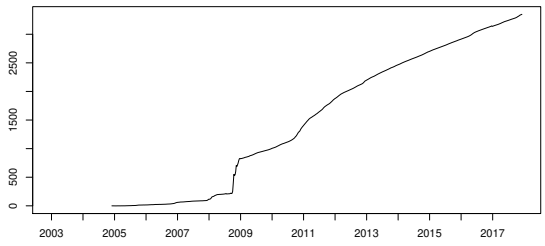


SV-APF/LLM-Bayesian

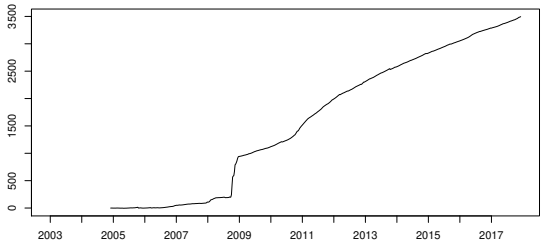
Figure 8: $LS_t^{\mathcal{M}}$ and LS_t^*, EUR-USD.



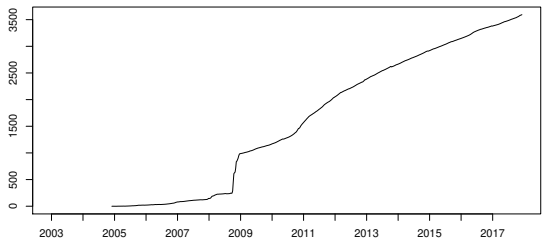
LLM-MLE



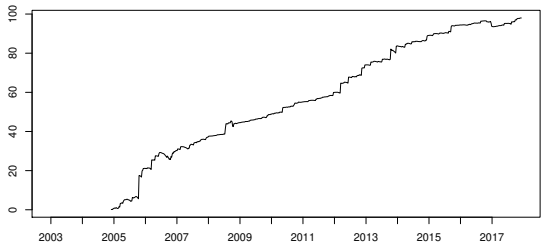
LLM-Bayesian



SV-FFBS



SV-APF



SV-APF/SV-FFBS

Figure 9: $LS_t^{\mathcal{M}}$ and LS_t^*, MXN-USD.

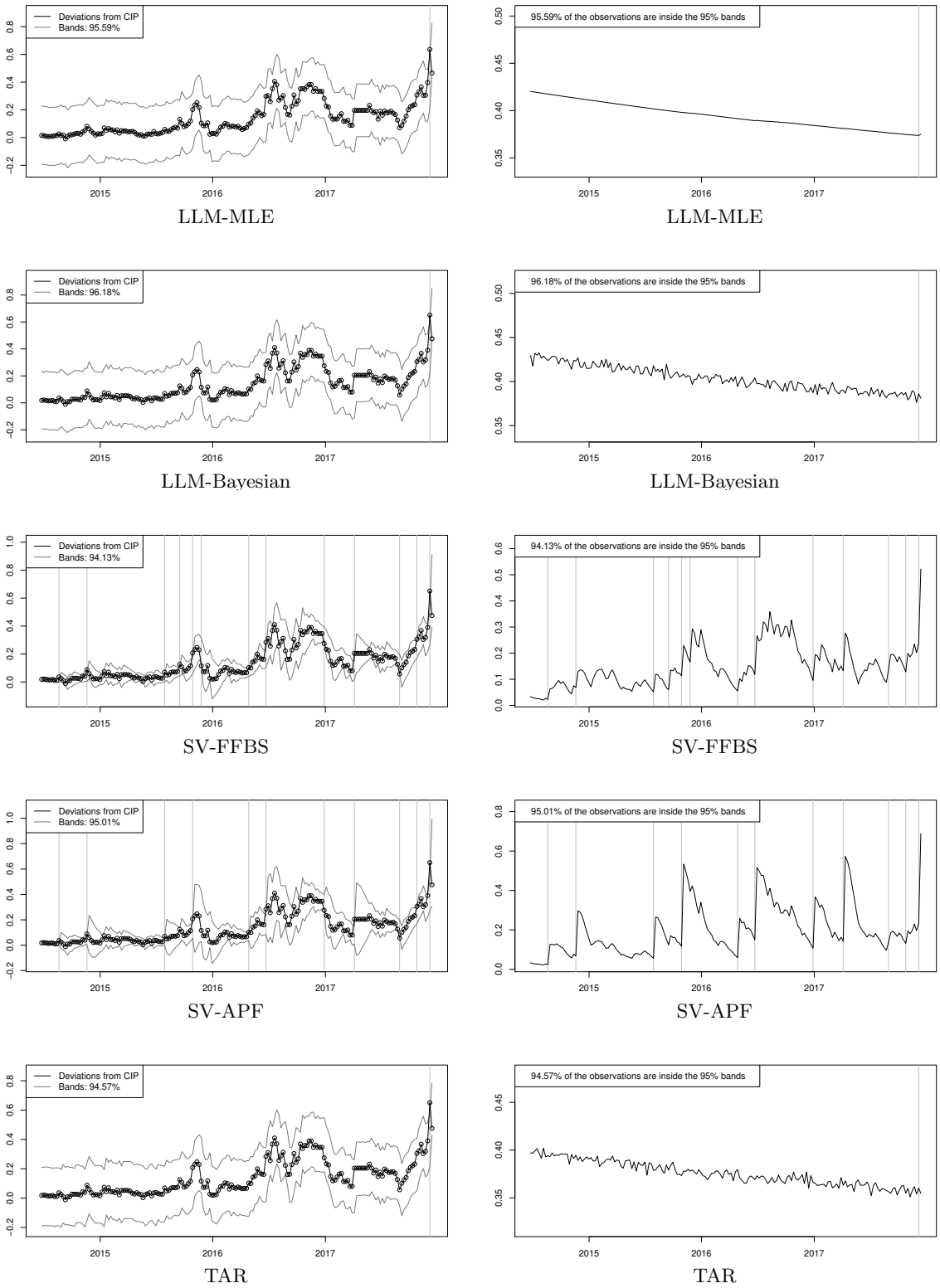


Figure 10: Deviations from CIP (left column) and estimated w in percentage points (right column) in percentage points, final 180 observations, GBP-USD. Vertical lines are observations outside w. Label contains the percentage of observations within w for the entire sample.

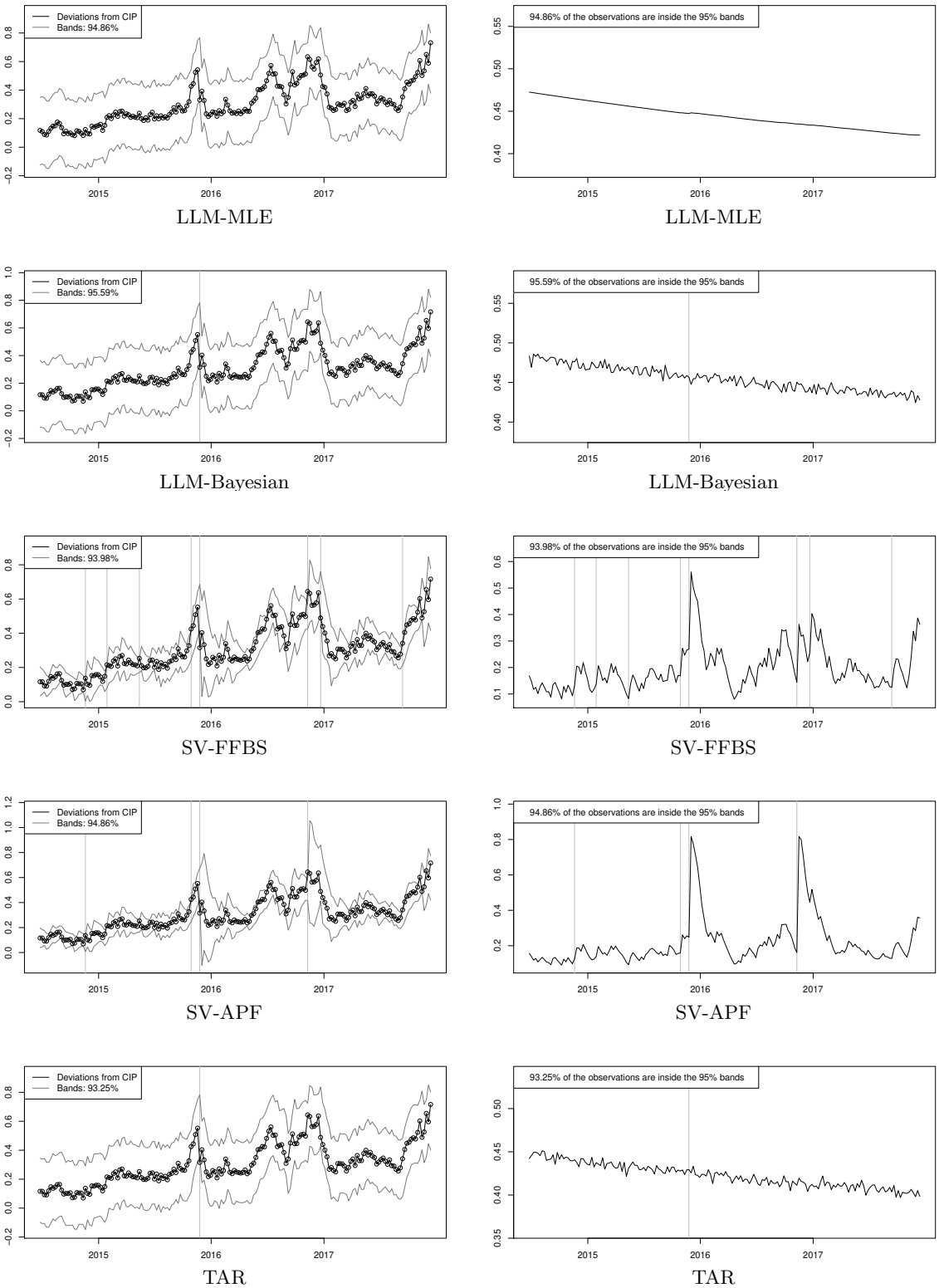


Figure 11: Deviations from CIP (left column) and estimated w in percentage points (right column) in percentage points, final 180 observations, EUR-USD. Vertical lines are observations outside w. Label contains the percentage of observations within w for the entire sample.

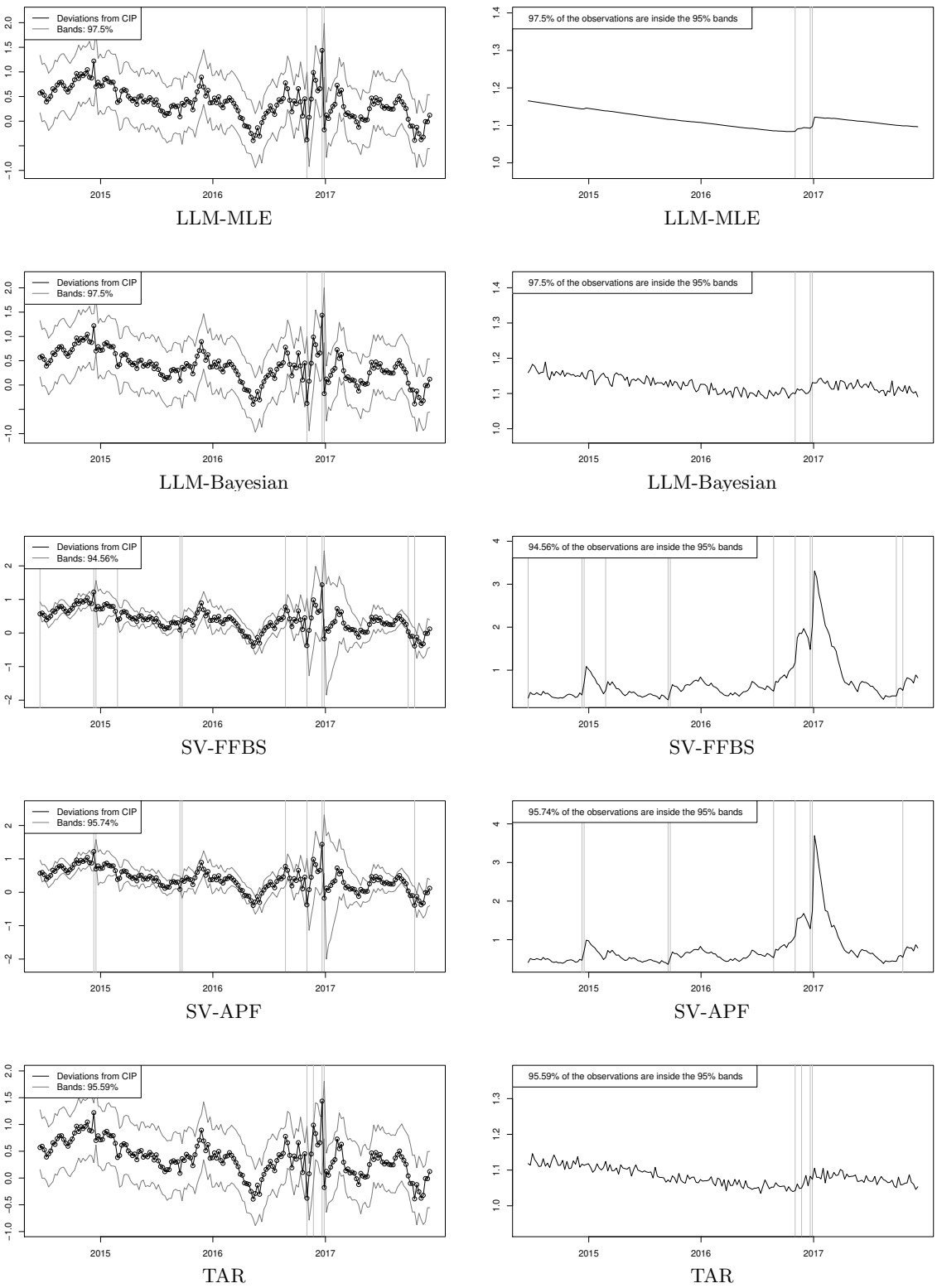


Figure 12: Deviations from CIP (left column) and estimated w in percentage points (right column) in percentage points, final 180 observations, MXN-USD. Vertical lines are observations outside w. Label contains the percentage of observations within w for the entire sample.

A Results for Sovereign Rate Differentials

A.1 Data

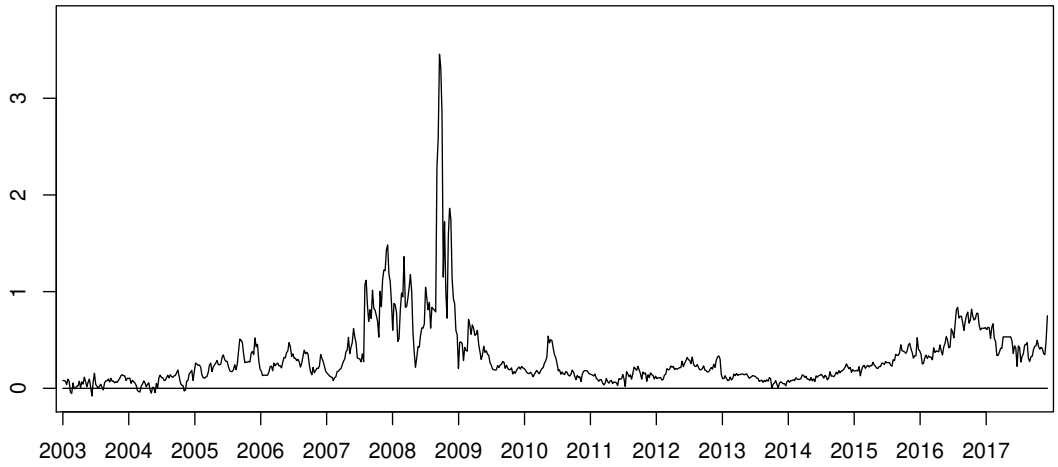


Figure 13: Deviations from CIP in percentage points GBP-USD. Computed with 3-month sovereign rates (Gilt-Treasuries). Sample: January 7th 2003 to December 26th 2017. Own calculations for weekly data. Source: Bank of England and Bloomberg.

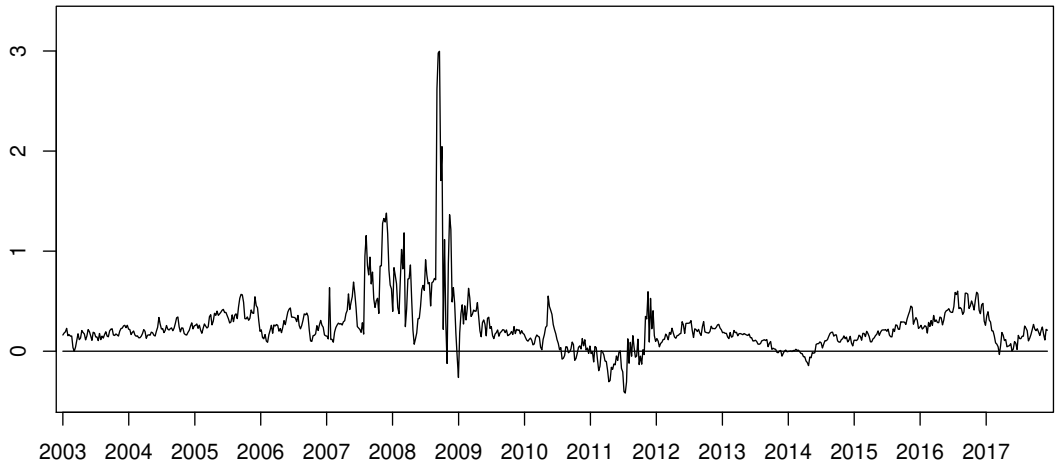


Figure 14: Deviations from CIP in percentage points EUR-USD. Computed with 3-month sovereign rates (German Bond-Treasuries). Sample: January 7th 2003 to December 26th 2017. Own calculations for weekly data. Source: Bloomberg.

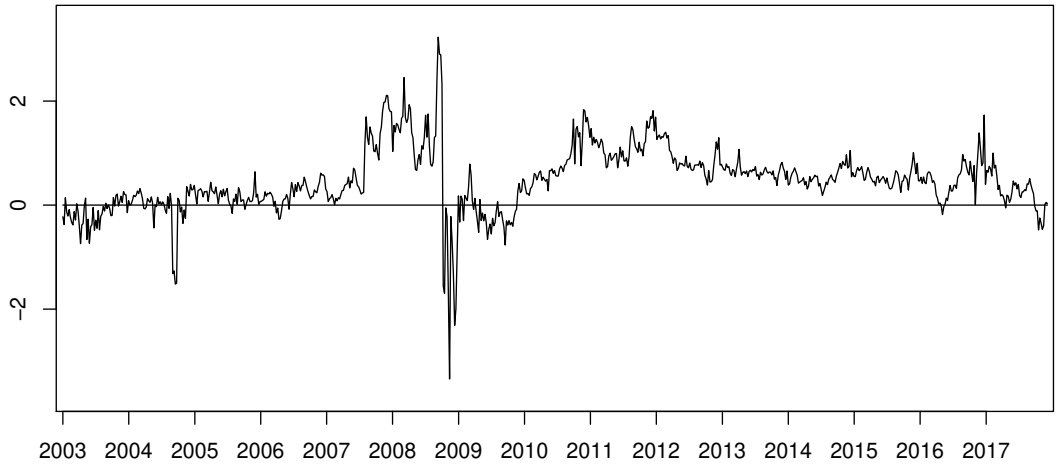


Figure 15: Deviations from CIP in percentage points MXN-USD. Computed with 3-month sovereign (CETES-Treasuries). Sample: January 7th 2003 to December 26th 2017. Own calculations for weekly data. Source: Banco de México and Bloomberg.

A.2 Width Estimates

Model/Period	Dec04-Jun07	Jun07-Dec09	Jan10-Dec17
	GBP		
LLM-MLE	18-21	20-78	52-73
LLM-Bayesian	19-23	22-79	52-74
SV-FFBS	14-41	13-388	7-55
SV-APF	17-29	14-404	8-99
TAR	15-20	20-77	49-71
	EUR		
LLM-MLE	16-28	28-94	64-88
LLM-Bayesian	17-29	29-95	64-89
SV-FFBS	16-91	12-472	6-145
SV-APF	16-43	15-346	8-162
TAR	15-26	26-90	60-83
	MXN		
LLM-MLE	90-122	88-168	122-160
LLM-Bayesian	89-122	88-171	121-161
SV-FFBS	30-191	38-899	26-336
SV-APF	39-187	50-712	35-266
TAR	82-113	82-157	112-150

Table 6: Estimated range for w in basis points using sovereign interest rate differentials.

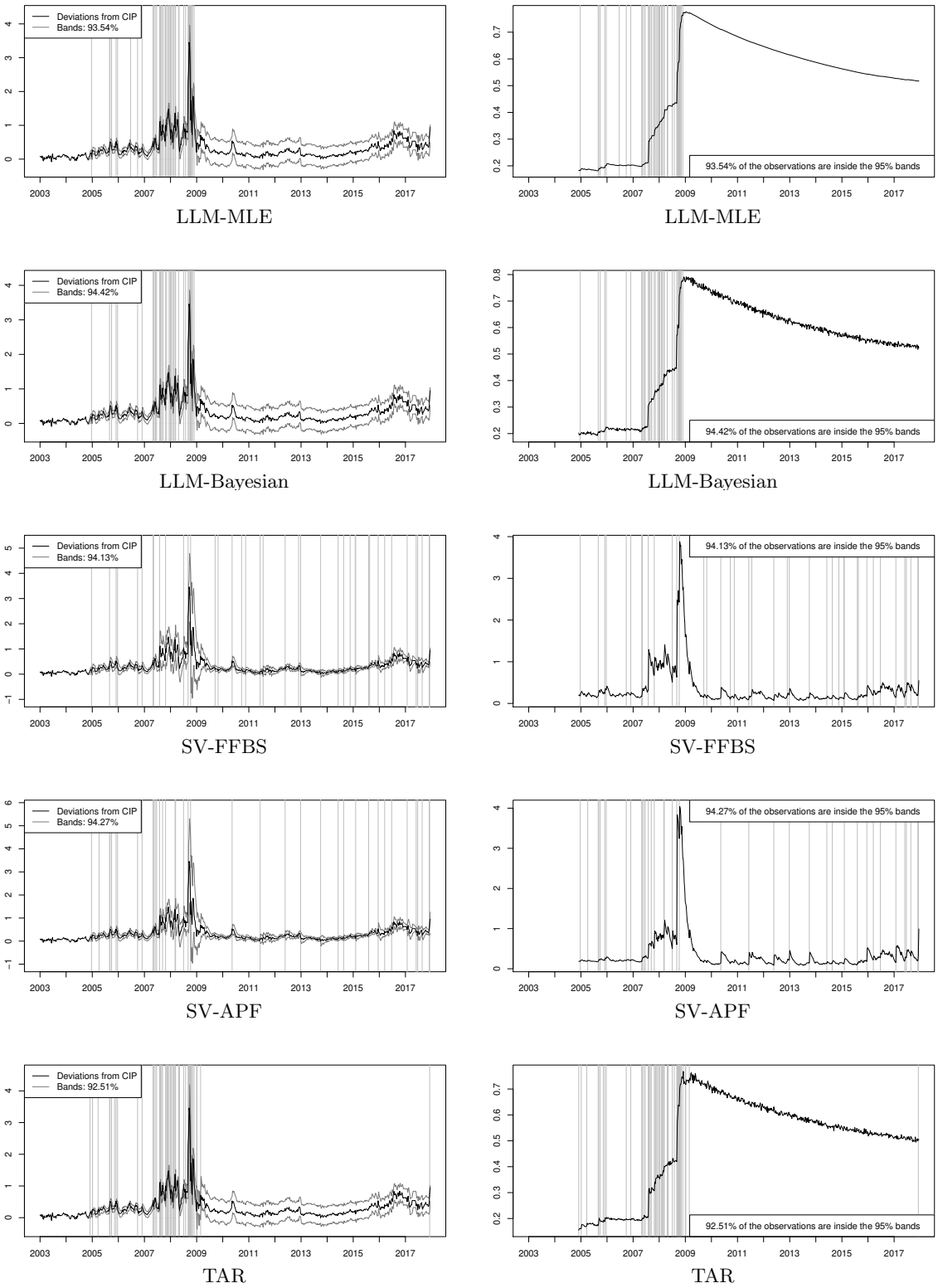


Figure 16: Deviations from CIP in percentage points and neutral band estimates GBP-USD computed with sovereign rates (left column). Estimated w in percentage points (right column). Vertical lines are observations outside w. Label contains the percentage of observations within w.

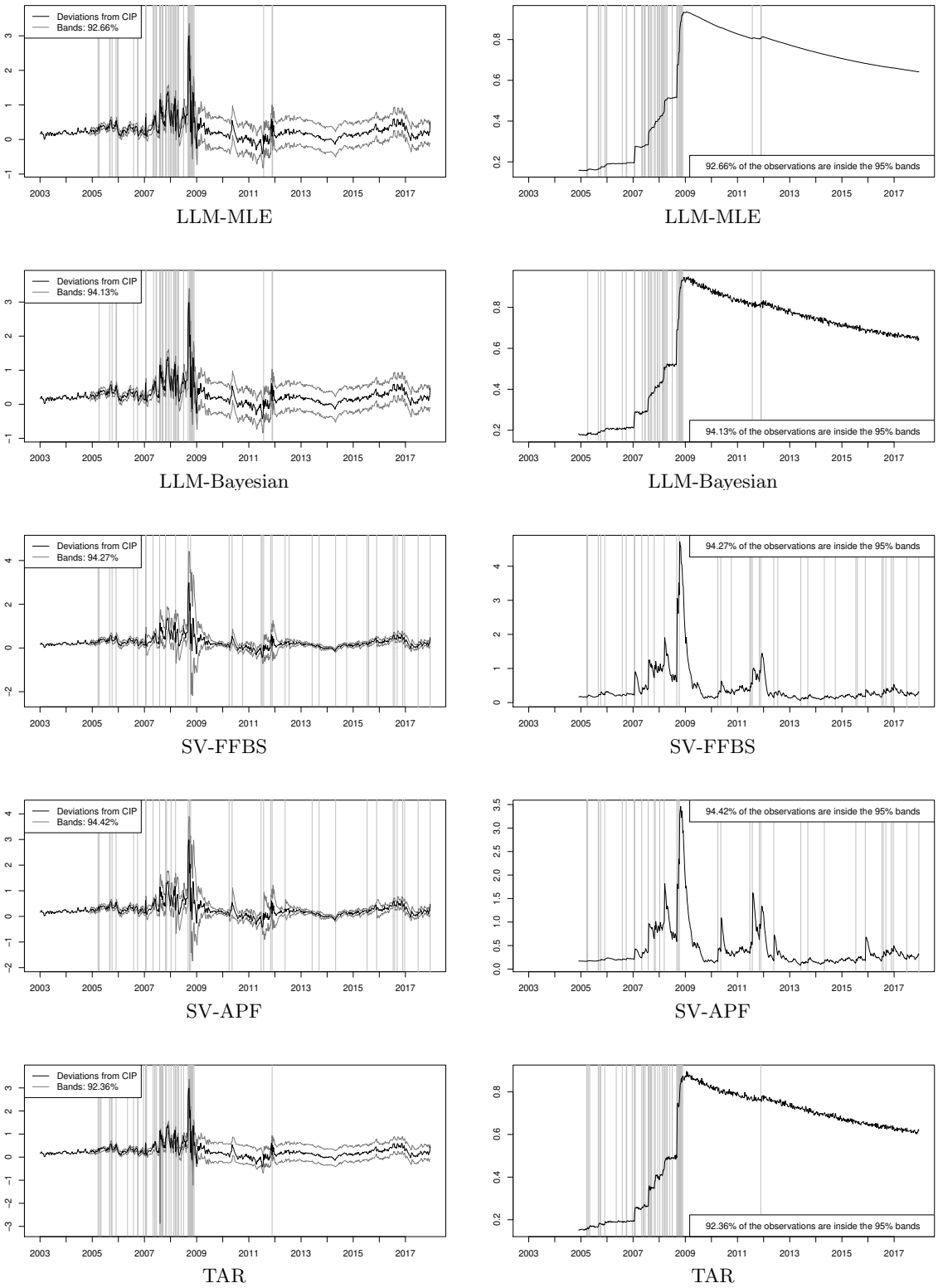


Figure 17: Deviations from CIP in percentage points and neutral band estimates EUR-USD computed with sovereign rates (left column). Estimated w in percentage points (right column). Vertical lines are observations outside w. Label contains the percentage of observations within w.

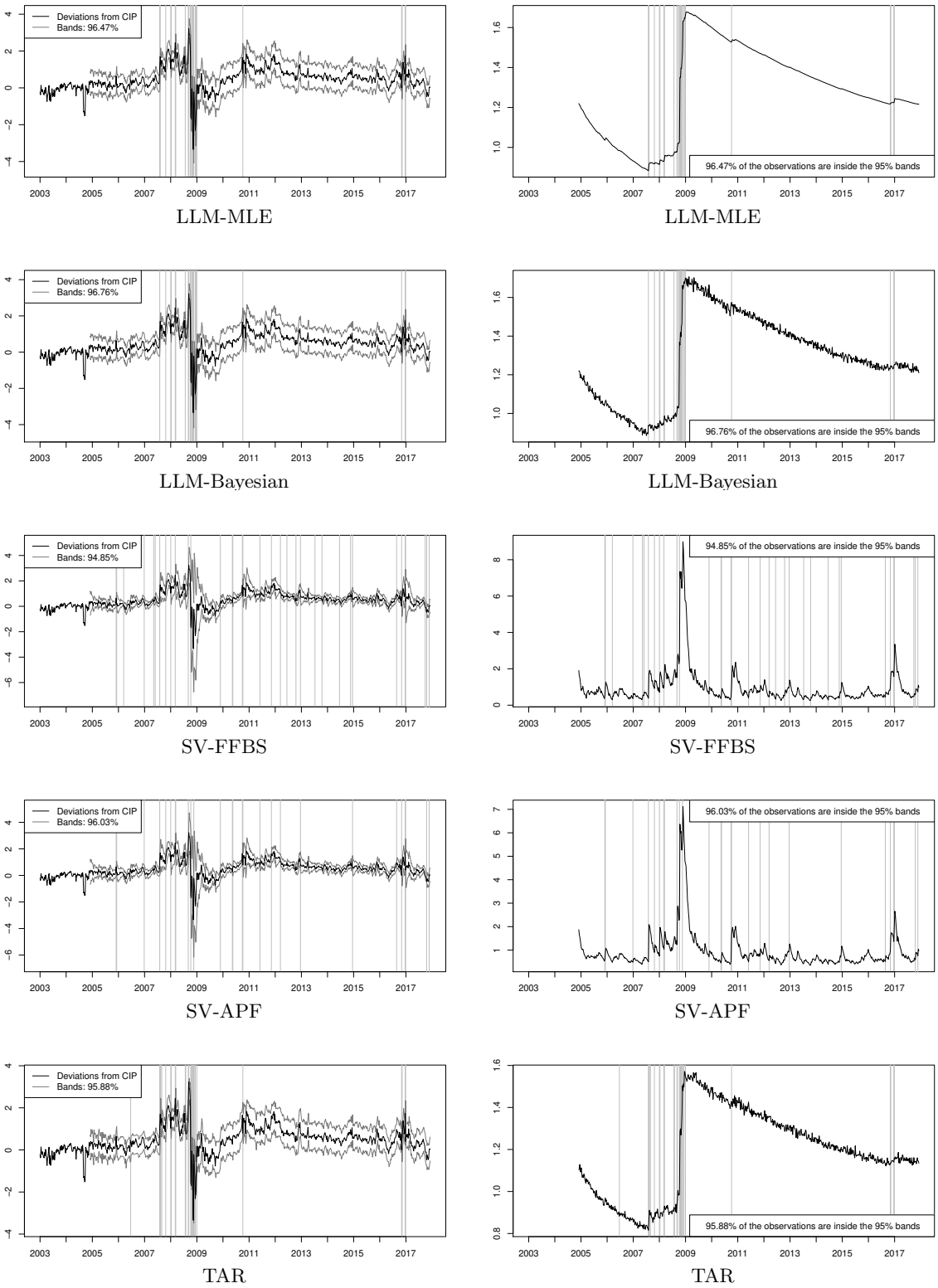


Figure 18: Deviations from CIP in percentage points and neutral band estimates MXN-USD computed with sovereign rates (left column). Estimated w in percentage points (right column). Vertical lines are observations outside w. Label contains the percentage of observations within w.

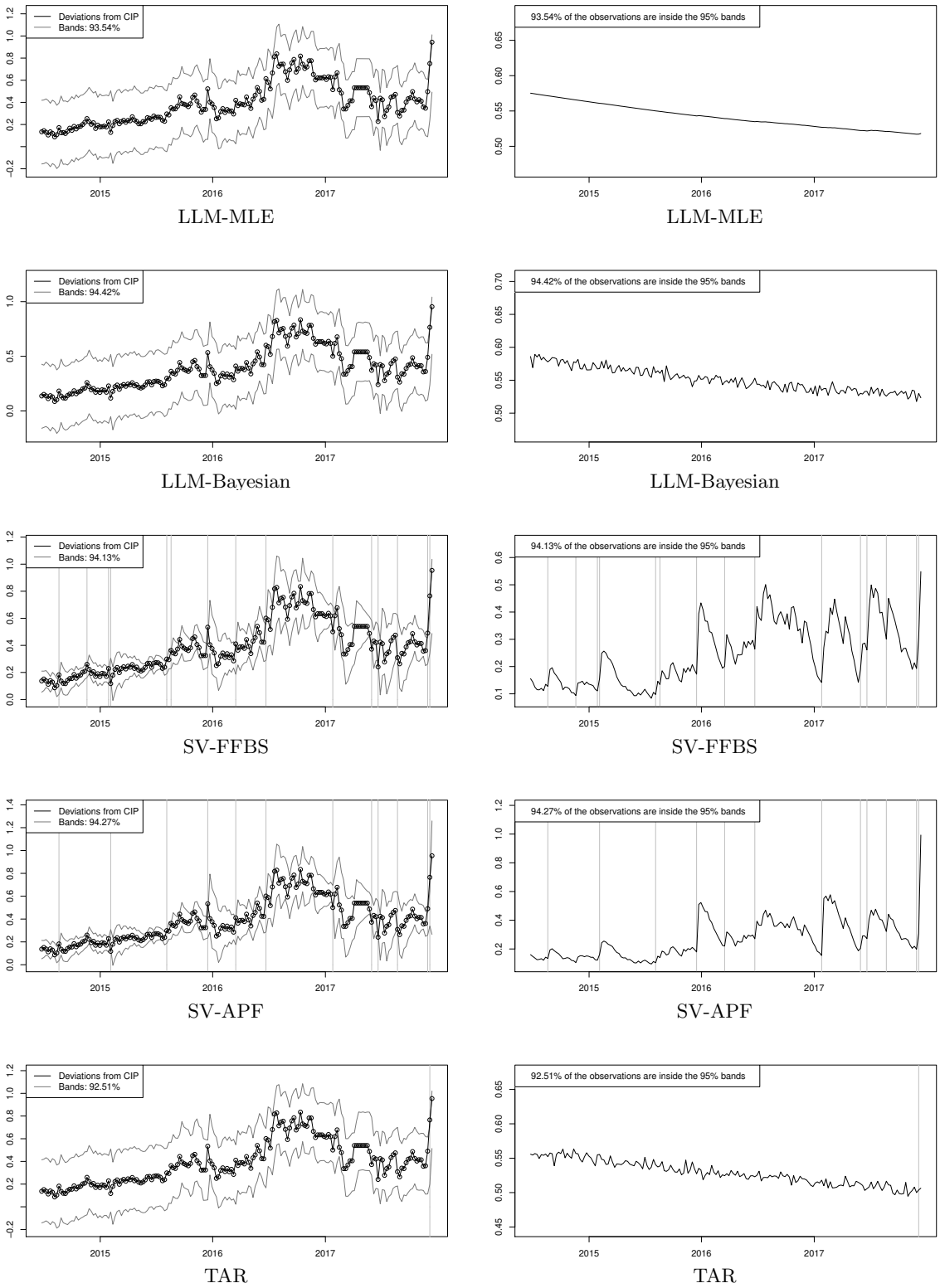


Figure 19: Deviations from CIP computed with sovereign rates (left column) and estimated w in percentage points (right column) in percentage points, final 180 observations, GBP-USD. Vertical lines are observations outside w. Label contains the percentage of observations within w for the entire sample.

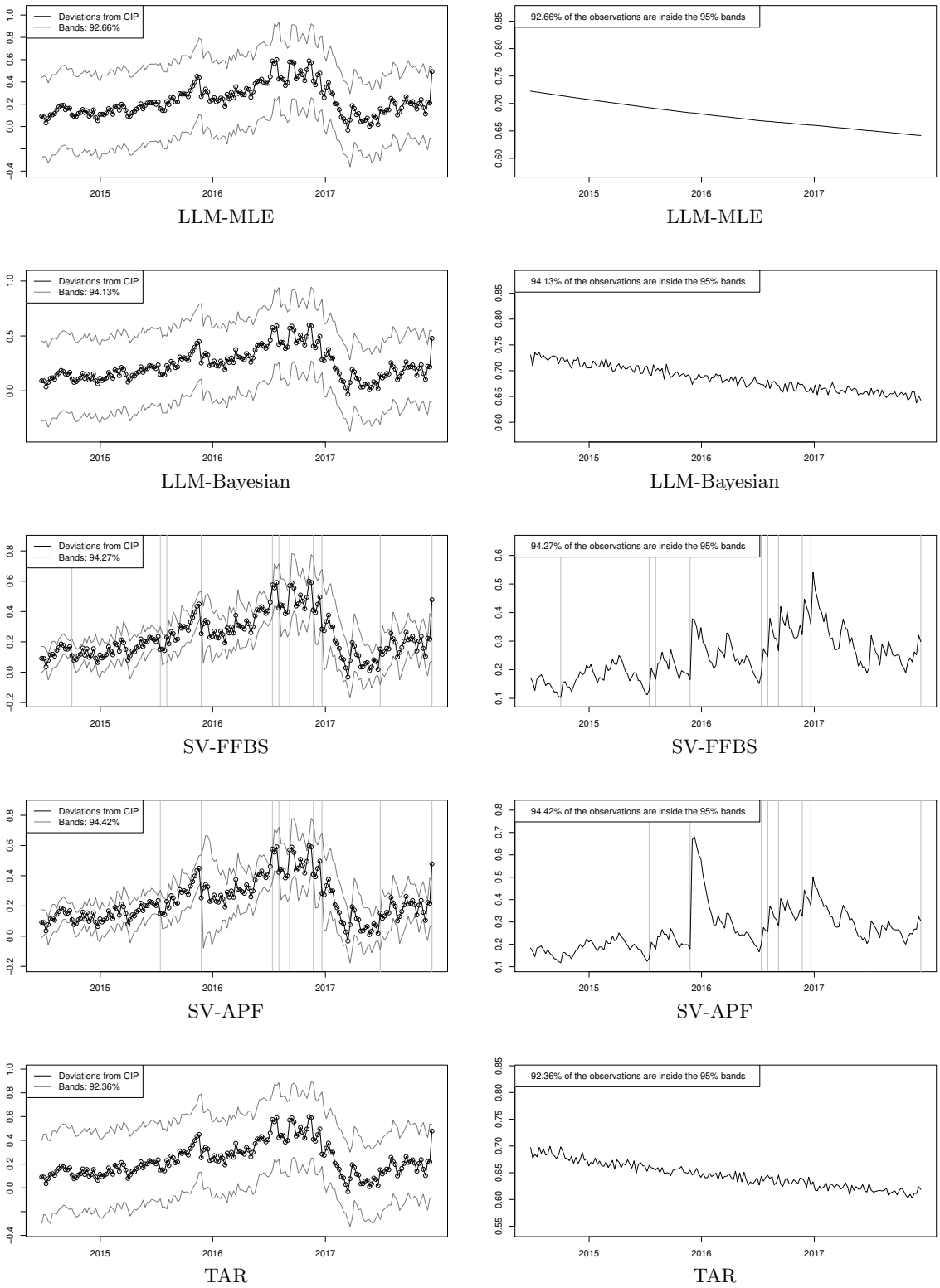


Figure 20: Deviations from CIP computed with sovereign rates (left column) and estimated w in percentage points (right column) in percentage points, final 180 observations, EUR-USD. Vertical lines are observations outside w. Label contains the percentage of observations within w for the entire sample.

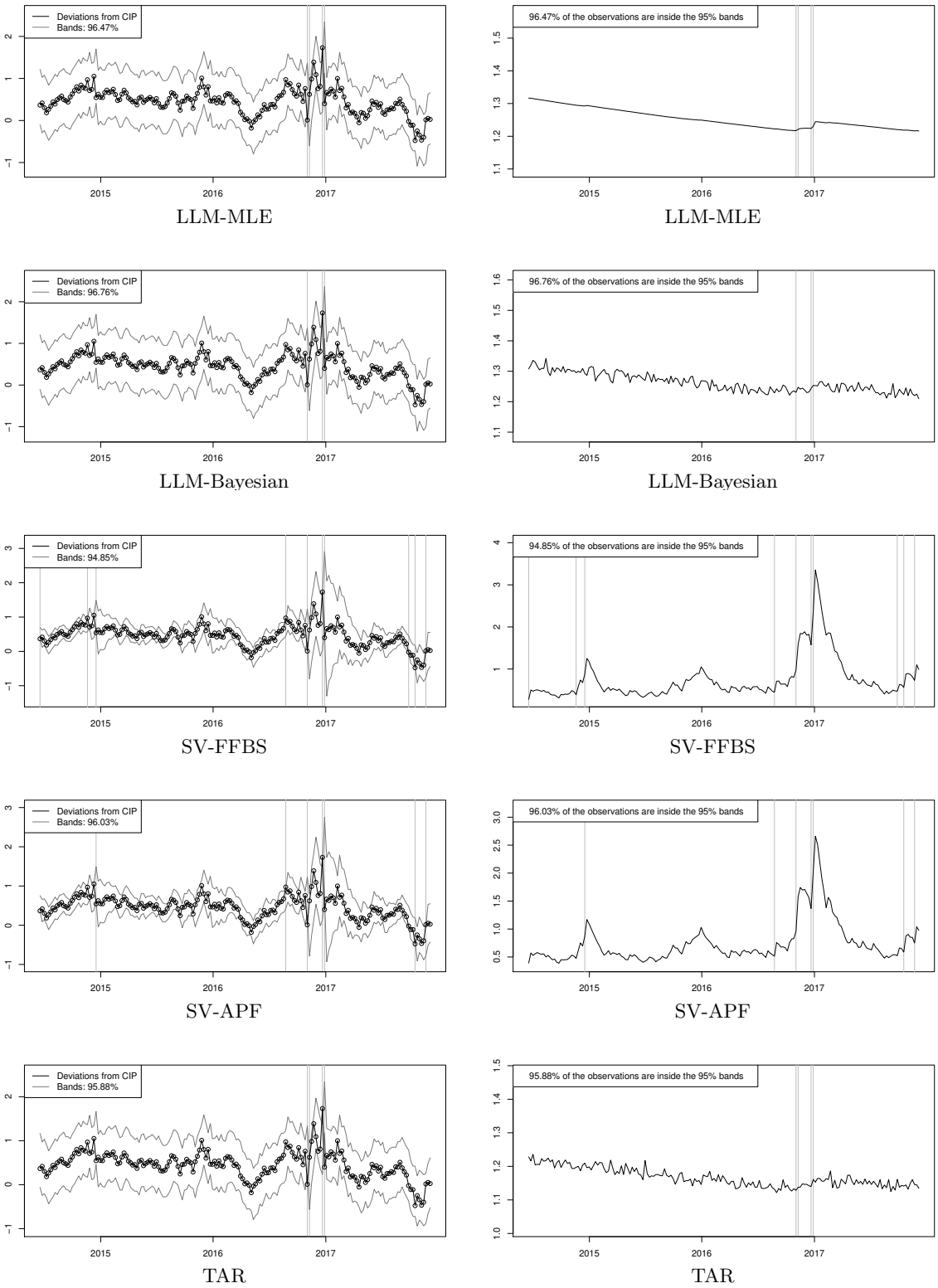


Figure 21: Deviations from CIP computed with sovereign rates (left column) and estimated w in percentage points (right column) in percentage points, final 180 observations, MXN-USD. Vertical lines are observations outside w. Label contains the percentage of observations within w for the entire sample.

A.3 Model Evaluation

	GBP-USD	EUR-USD	MXN-USD
Sovereign			
LLM-MLE	-104.6278	-281.2509	-542.4628
LLM-Bayesian	48.6477	-117.5548	-480.2678+
SV-FFBS	164.3163+	-14.71815+	-554.4423
SV-APF	289.9309*	85.2192*	-326.1555*
TAR	-4484.403	-2647.52	-2795.97

Table 7: $CLPL_n$. * Refers to the combination with highest likelihood. + Is the second highest.

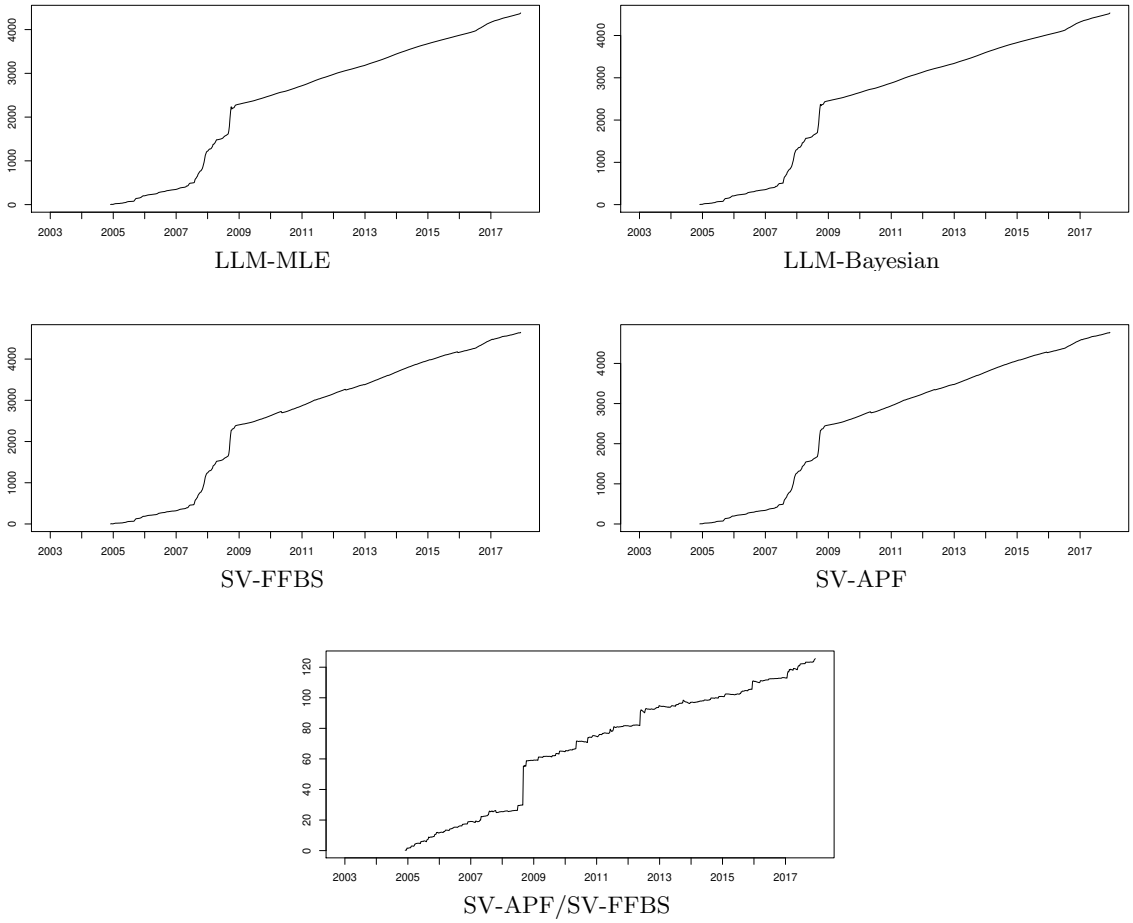
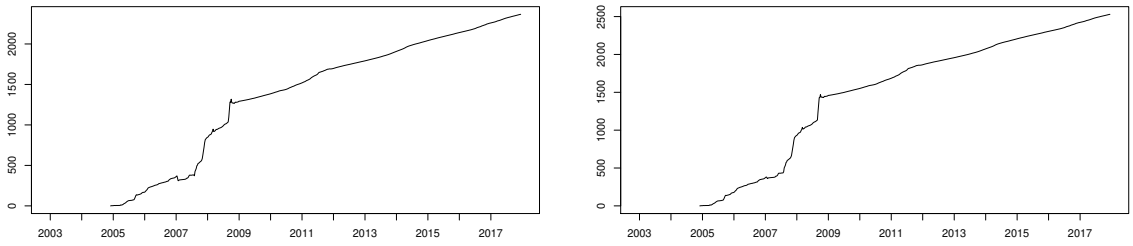
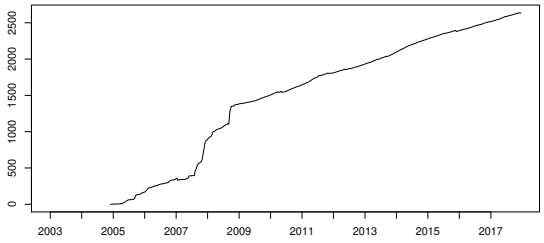


Figure 22: $LS_t^{\mathcal{M}}$ and LS_t^*, GBP-USD. Computed with sovereign rates.

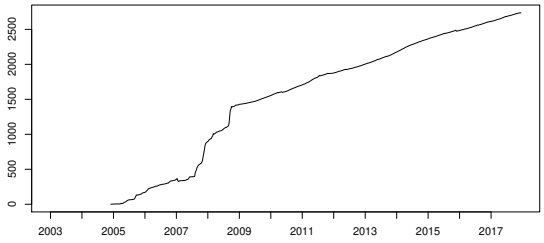


LLM-MLE

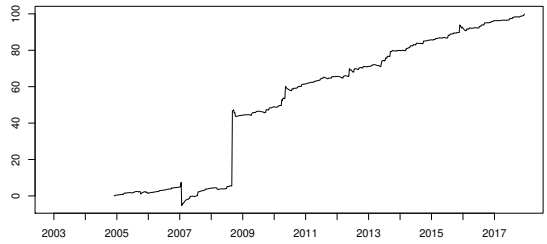
LLM-Bayesian



SV-FFBS

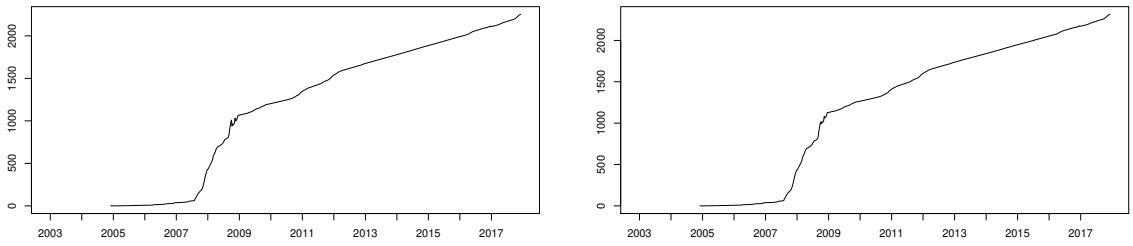


SV-APF



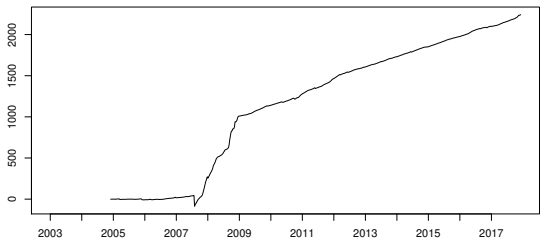
SV-APF/SV-FFBS

Figure 23: $LS_t^{\mathcal{M}}$ and LS_t^*, EUR-USD. Computed with sovereign rates.

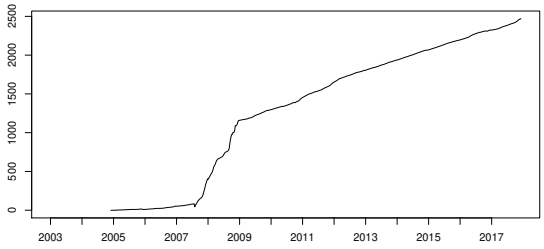


LLM-MLE

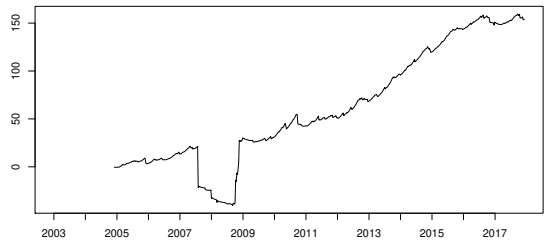
LLM-Bayesian



SV-FFBS



SV-APF



SV-APF/LLM-Bayesian

Figure 24: $LS_t^{\mathcal{M}}$ and LS_t^*, MXN-USD. Computed with sovereign rates.