Received: 17 December 2020 | Revised: 26 February 2023 | Accepted: 26 March 2023

DOI: 10.1002/jae.2984

WILEY Journal of APPLIED ECONOMETRICS

RESEARCH ARTICLE

Exchange rates and macroeconomic fundamentals: Evidence of instabilities from time-varying factor loadings

Eric Hillebrand[1] | **Jakob Guldbæk Mikkelsen**[2] | **Lars Spreng**[3] | **Giovanni Urga**[3]

[1]Department of Economics and Business Economics, Aarhus University, Aarhus, Denmark

[2]Danmarks Nationalbank, Copenhagen, Denmark

[3]Centre for Econometric Analysis and Faculty of Finance, Bayes Business School (formerly Cass), London, UK

Correspondence
Eric Hillebrand, Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark.
Email: ehillebrand@econ.au.dk

Summary

We examine the relationship between exchange rates and macroeconomic fundamentals using a two-step maximum likelihood estimator through which we compute time-varying factor loadings. Factors are obtained as principal components, extracted from vintage macro-datasets that combine FRED-MD and OECD databases. Using 14 currencies over 1990–2021, we show that the loadings on the factors vary considerably over time and increase the percentage of explained variation in exchange rates by an order of magnitude. Time-varying loadings improve the overall predictive ability of the model, especially during crises, and lead to better forecasts of sign changes in exchange rates.

KEYWORDS

exchange rate forecasting, foreign exchange rates, high-dimensional factor models, macroeconomic factors, time-varying loadings

1 | INTRODUCTION

In this paper, we analyse the unstable relationship between exchange rates and macroeconomic fundamentals. To this end, we apply the two-step maximum likelihood approach proposed in Mikkelsen et al. (2019) that enables the estimation of time-varying loadings (TVL) in factor models.

Two-step estimation in large factor models was proposed in Doz et al. (2011, 2012). These, together with the important result found in Bates et al. (2013) that principal components are consistent estimators of unobserved factors even in the presence of time-varying loadings, allowed Mikkelsen et al. (2019) to propose the consistent estimation of time-varying factor loadings in two-step maximum likelihood. Different approaches are taken in Su and Wang (2017), who estimate smoothly changing time-varying factor loadings using a local principal component estimator for latent factors, and Barigozzi et al. (2021), who introduce a generalised dynamic factor model in which factors are loaded with a time-varying filter.

Since Meese and Rogoff's (1983) key finding that structural exchange rate models perform no better than a random walk, arduous empirical work has been invested into this *disconnect puzzle* (Obstfeld & Rogoff, 2000), however, in many cases to no avail (Rossi, 2013). One possible solution for the disconnect puzzle is to model the relationship between macroeconomic fundamentals and exchange rates as time-varying. A theoretical explanation for the presence of time-variation is provided by the scapegoat theory (Bacchetta & van Wincoop, 2004, 2012, 2013). An observed fundamental becomes a scapegoat if it is correlated with an unobserved shock and investors attribute exchange rate fluctuations to the fundamental instead of the actual unobserved shock. This leads investors to update their expectation about the effect fundamentals exert in a time-varying manner. Fratzscher et al. (2015) are the first to present empirical support for this theory by examining a survey of FX traders to obtain a measure of the scapegoat weights. In the same vein, Pozzi and Sadaba (2020) construct parameter expectations from survey data and use a Bayesian approach to determine the probability that variables are scapegoats. The disconnect puzzle may, however, also be a product of inaccurate model selection, as suggested by Sarno and Valente (2009): Models would have to be altered frequently to optimally capture the information embedded in fundamentals, and this implies a high degree of time-variation in their parameters. Kouwenberg et al. (2017) develop a dynamic model selection rule, which they find to produce better forecasts than several benchmark models. The reason behind this lies, again, in the rule's ability to incorporate time-variation. Further evidence for parameter instability in exchange rate regressions is provided by Rossi (2006), Bekiros (2014), and Byrne et al. (2018).

The present paper uses the results of Mikkelsen et al. (2019) to estimate a theoretical model in which the relationship between exchange rates and macroeconomic fundamentals is unstable. We treat macroeconomic fundamentals as latent factors, which are extracted in real-time from 272 newly compiled monthly vintage datasets. Specifically, we collate all real-time vintages of the McCracken and Ng (2016) FRED-MD database from 1999:08 to 2022:02. In addition, we compile large vintage datasets from the OECD statistical database for the same time periods, which we merge with the FRED-MD data. This yields a novel real-time database with each series starting in 1990:04 and ending 1 month prior to their release.[1] Therefore, rather than just relying on first releases, we include revisions made to past data with each new release. That is, we accurately replicate the information set available at each time step, which has been shown to improve forecasts of financial variables (Caruso & Coroneo, 2022). The information inherent in these series is extracted via principal components that serve as factor estimates. The factors can be interpreted as real economy, housing market, and interest rate factors. The model is tested for 14 different currencies vis-à-vis the US dollar. To examine whether accounting for time variation improves exchange rate predictions, we compare the results to a factor model with constant loadings. As constant factor models have been found to deliver limited to no forecast improvements over a random walk (Engel et al., 2015; Rossi, 2013), we also analyse whether incorporating time-variation increases predictive ability relative to a random walk. The paper provides in-sample evidence that accounting for time-variation improves the model fit considerably as demonstrated by an R^2 ranging from 27% up to 88%. It correctly matches an appreciation and depreciation up to 89% of the time, whereas the constant loadings model can only explain a very small part of exchange rate variations, demonstrated by an R^2 between 0% and 6%. We show that taking the aforementioned instabilities into consideration improves the out-of-sample forecast accuracy of the model across benchmarks and evaluation procedures. First, the time-varying model displays better predictive ability than the constant loadings model across different forecasting horizons according to the conditional and unconditional Giacomini and White (2006) test. The constant loadings model never outperforms the time-varying model. Second, the time-varying model's forecasts display greater direction accuracy. Third, it performs better than the constant loadings model relative to a random walk. It achieves a lower root mean squared error (RMSE) than the random walk for up to nine currencies and outperforms it statistically in two cases. In line with Engel et al. (2015), we find that the constant model can achieve a lower RMSE only at very long forecasting horizons. Fourth, when comparing predictive ability locally using the Giacomini and Rossi (2010) fluctuation test, we observe that the time-varying model improves forecasts during crises and, for several exchange rates, performs well against the random walk during periods in which the constant model displays poor performance.

The paper is structured as follows: Section 2 presents the theoretical model of structural instabilities between exchange rates and fundamentals. In Section 3, the model is mapped into state space form, and the econometric approach is described. Section 4 discusses the data; in Section 5, we report the in-sample; and in Section 6, the out-of-sample results. Section 7 concludes.

2 | THEORETICAL MODEL

This section derives a model with instabilities in the relationship between exchange rates and macroeconomic fundamentals. The model belongs to the same class as the ones examined by Engel and West (2005). That is, it expresses the exchange rate as the discounted value of expected future fundamentals and unobservable shocks (see equation (1) in

[1]That is, the 1999:08 vintage dataset starts in 1990:04 and ends in 1999:07. The 1999:09 vintage starts in 1990:04 and ends in 1999:08, and so on.

Engel & West, 2005). Specifically, the relation is

$$\Delta s_t = F_t + \sum_{j=1}^{\infty} \left(\frac{1}{\mu}\right)^j \mathbb{E}[F_{t+j}|I_t] - \sum_{j=0}^{\infty} \left(\frac{1}{\mu}\right)^{j+1} \mathbb{E}[\phi_{t+j}|I_t], \tag{1}$$

where s_t is the log of the exchange rate measured as the domestic price per unit of foreign currency, $\mathbb{E}[\cdot|I_t]$ is the expectation of the representative agent conditional on I_t, the information set available at time t , and $\mu \geq 0$. The value of the exchange rate is determined by the present and expected future macroeconomic fundamentals F_t. Finally, ϕ_t is the risk premium. The above equation results from two conditions. The first is an uncovered interest parity (UIP) condition[2]:

$$\mathbb{E}[s_{t+1}|I_t] - s_t = i_t - i_t^* + \phi_t, \tag{2}$$

where i_t is the nominal one-period interest rate. An asterisk denotes foreign variables, and deviations from UIP are accounted for by the risk premium ϕ_t. The expected change in the exchange rate is thus equal to the interest rate differential between the domestic and the foreign country plus a risk premium. The second condition relates the interest rate differential to macroeconomic fundamentals:

$$i_t - i_t^* = \mu \Delta s_t - \mu F_t. \tag{3}$$

Engel and West (2005) discuss a range of models that lead to Equations (1) and (3), for instance, a Taylor rule or monetary model, and it is also derived in Bacchetta and van Wincoop (2013, equation 3).[3] Combining Equations (2) and (3) results in

$$\begin{aligned} \mathbb{E}[\Delta s_{t+1}|I_t] &= \mu \Delta s_t - \mu F_t + \phi_t \\ \Delta s_t &= \frac{1}{\mu} \{\mathbb{E}[\Delta s_{t+1}|I_t] - \phi_t\} + F_t. \end{aligned}$$

Recursive substitution of Δs_t, assuming no bubbles, yields Equation (1), which establishes the common result that the exchange rate equals the present value of expected future macroeconomic fundamentals and the foreign exchange risk premium. In their model, Bacchetta and van Wincoop (2013) specify $F_t = f_t'\beta$ as a linear combination of observable macroeconomic fundamentals, where $f_t = (f_{1,t}, \ldots f_{n,t})'$ and $\beta = (\beta_1, \cdots, \beta_n)'$. We allow this combination to be time-varying:

$$F_t = f_t'\xi_t = f_t'(\beta + \kappa_t).$$

That is, consistent with the theory in Bacchetta and van Wincoop (2013), β describes a long-run equilibrium relationship between fundamentals and exchange rates, but we allow for transitory, zero-mean deviations in form of $\kappa_t = (\kappa_{1,t}, \cdots, \kappa_{n,t})'$. Consequently, the relative importance of fundamentals in determining the exchange rate is time-varying and affected by κ_t. This specification nests the constant coefficients case if $\kappa_t = 0$ for all t. To derive the effect of changes in observed fundamentals on the exchange rate, consider for simplicity the case of a single fundamental and assume that f_t, κ_t, and ϕ_t follow AR(1) processes:

$$\begin{aligned} f_t &= \rho_f f_{t-1} + v_t, & v_t &\sim i.i.d.(0, \sigma_f^2) \\ \kappa_t &= \rho_\kappa \kappa_{t-1} + u_t, & u_t &\sim i.i.d.(0, \sigma_\kappa^2) \\ \phi_t &= \rho_\phi \phi_{t-1} + w_t, & w_t &\sim i.i.d.(0, \sigma_\phi^2), \end{aligned} \tag{4}$$

[2]See Engel (2014) for a survey of exchange rates and interest parity as well as the existence of the risk premium term.

[3]Note that Bacchetta and van Wincoop (2013) specify their model in levels and distinguish between observed fundamentals (F_t) and unobserved fundamentals (b_t). They obtain $i_t - i_t^* = \mu s_t - \mu(F_t + b_t)$ which they show to be consistent with several established exchange rate models, such as the monetary model. In an earlier version, they show that this relationship also holds in first differences (Bacchetta & van Wincoop, 2009).

where $|\rho_f|, |\rho_\kappa|, |\rho_\phi| < 1$. Clearly, $\mathbb{E}[f_{t+j}|I_t] = \rho_f^j f_t$ and $\mathbb{E}[\kappa_{t+j}|I_t] = \rho_\kappa^j \kappa_t$. Assuming f_t and κ_t are uncorrelated, Equation (1) becomes:

$$\begin{aligned}\Delta s_t &= \sum_{j=0}^{\infty}\left(\frac{1}{\mu}\right)^j \rho_f^j f_t\beta + \sum_{j=0}^{\infty}\left(\frac{1}{\mu}\right)^j \rho_\kappa^j f_t\kappa_t - \frac{1}{\mu}\sum_{j=0}^{\infty}\left(\frac{1}{\mu}\right)^j \rho_\phi^j \phi_t \\ &= f_t\left(\frac{\mu}{\mu-\rho_f}\beta + \frac{\mu}{\mu-\rho_\kappa}\kappa_t\right) - \frac{1}{\mu-\rho_\phi}\phi_t.\end{aligned} \tag{5}$$

The derivative of the exchange rate with respect to the fundamentals is

$$\frac{\partial \Delta s_t}{\partial f_t} = \left(\frac{\mu}{\mu-\rho_f}\beta + \frac{\mu}{\mu-\rho_\kappa}\kappa_t\right). \tag{6}$$

That is, the effect of variations in macroeconomic fundamentals on the exchange rate corresponds to a constant part, $\frac{\mu}{\mu-\rho_f}\beta$, and a time-varying part, $\frac{\mu}{\mu-\rho_\kappa}\kappa_t$. Based on existing studies, one can expect that μ is close to one (Engel & West, 2005). In that case, if the transitory shocks are highly persistent relative to the fundamentals, the relationship between the latter and exchange rates is characterised by a greater degree of instability.

3 | ECONOMETRIC ESTIMATION

3.1 | State space formulation

This subsection demonstrates how the theoretical model can be mapped into state space form. Focus on the single factor case for illustrative purposes and combine Equation (5) with the autoregressive processes for the transitory shocks to obtain the system:

$$\begin{aligned}\kappa_t &= \rho_\kappa\kappa_{t-1} + u_t, \\ \Delta s_t &= f_t\left(\frac{\mu}{\mu-\rho_f}\beta + \frac{\mu}{\mu-\rho_\kappa}\kappa_t\right) - \frac{1}{\mu-\rho_\kappa}\phi_t.\end{aligned} \tag{7}$$

To estimate the relation of exchange rates to fundamentals, $\frac{\mu}{\mu-\rho_f}\beta + \frac{\mu}{\mu-\rho_\kappa}\kappa_t$, write the system in the following state space representation:

$$\begin{aligned}\lambda_t - \bar{\lambda} &= b(\lambda_{t-1} - \bar{\lambda}) + \eta_t, \qquad & \eta_t &\sim i.i.d.(0, \sigma_\eta^2), \\ \Delta s_t &= f_t'\lambda_t + \varepsilon_t, & \varepsilon_t &\sim i.i.d.(0, \sigma_\varepsilon^2),\end{aligned} \tag{8}$$

where the measurement error ε_t is an estimate of the risk premium[4] and the state vector λ_t estimates the relation between macroeconomic fundamentals f_t and the exchange rate s_t. By comparing Equations (7) and (8), it can be seen that $\lambda_t = \frac{\mu}{\mu-\rho_f}\beta + \frac{\mu}{\mu-\rho_\kappa}\kappa_t$; hence, the parameters of the state space representation (8) can be mapped to the parameters in Equation (7):

$$\begin{aligned}\bar{\lambda} = \mathbb{E}[\lambda_t] &= \mathbb{E}\left[\frac{\mu}{\mu-\rho_f}\beta + \frac{\mu}{\mu-\rho_\kappa}\kappa_t\right] = \frac{\mu}{\mu-\rho_f}\beta, \\ \frac{\sigma_\eta^2}{1-b^2} = \mathbb{V}[\lambda_t] &= \mathbb{V}\left[\frac{\mu}{\mu-\rho_f}\beta + \frac{\mu}{\mu-\rho_\kappa}\kappa_t\right] = \mathbb{V}[\kappa_t] = \omega^2\frac{\sigma_u^2}{1-\rho_\kappa^2},\end{aligned}$$

where $\omega = \frac{\mu}{\mu-\rho_\kappa}$ and $\mathbb{V}$ denotes the variance. The autocorrelation parameter ρ_κ of κ_t corresponds to the autocorrelation parameter b of λ_t. Therefore, estimating state space system (8) will give estimates of the parameter vector $\frac{\mu}{\mu-\rho_f}\beta$ and estimates of the transitory shock process scaled by ω. The state space representation (8) can easily be generalised to the multivariate case with r observed fundamentals $f_t = (f_{1t}, \cdots, f_{rt})'$ and state vector $\lambda_t = (\lambda_{1t}, \cdots, \lambda_{rt})'$:

$$\begin{aligned}B(L)(\lambda_t - \bar{\lambda}) &= \eta_t, \qquad & \eta_t &\sim i.i.d.(0, Q), \\ \Delta s_t &= f_t'\lambda_t + \varepsilon_t, & \varepsilon_t &\sim i.i.d.(0, \sigma_\varepsilon^2),\end{aligned} \tag{9}$$

[4]Specifically, we have $\varepsilon = -\frac{1}{\mu-\rho_\phi}\phi_t$ and $\sigma_\varepsilon^2 = \frac{\sigma_\phi^2}{(\mu-\rho_\phi)^2(1-\rho_\phi^2)}$.

where $B(L) = I - B^0_{t,1}L - \cdots - B^0_{t,q}L^q$ is a q^{th}-order lag polynomial with roots outside the unit circle. The covariance matrix of the state innovation, η_t, is $\mathbb{E}[\eta_t\eta_t'] = Q$.

3.2 | A factor model with time-varying loadings

To estimate the system (9) empirically, we specify a factor model with time-varying loadings. We use a large panel of macroeconomic data series $X_t = (X_{1t}, \cdots, X_{Nt})'$, $t = 1, \cdots, T$, whereby we assume that X_{it} has a factor structure:

$$X_{it} = \alpha_{it}' f_t + \epsilon_{it},$$

where f_t is an $r \times 1$ vector of common factors, α_{it} are the corresponding time-varying factor loadings, and ϵ_{it} are idiosyncratic errors. In our application below, N and T are of the same order of magnitude, which renders one-step maximum likelihood estimation of the model infeasible due to the number of parameters to be estimated (Bai, 2003; Shumway & Stoffer, 1982).

Progress towards estimability of factor models by maximum likelihood has been made in Doz et al. (2011, 2012), who established a two-step procedure that first uses principal components to estimate the factors. To identify the effect of fundamentals on the exchange rate in the presence of structural instabilities in this paper, we employ two important theoretical results: (i) The principal component estimator gives consistent factor estimates even in the presence of time-varying loadings (Bates et al., 2013). (ii) Maximising the likelihood of a factor model with principal components as estimators of the unobservable factors gives consistent estimates of stationary time-varying loadings (Mikkelsen et al., 2019). We refer to the latter paper for details on how the estimation error from the first step is controlled and consistency is established.

The factor model allows the idiosyncratic errors to have limited cross-sectional correlation. The number of factors, r, is considerably smaller than the number of series, N, such that the information in the large number of macroeconomic variables is condensed into the r-dimensional factors. That is, by extracting the first r principal components of X_t, one can construct a set of macroeconomic factors that represents the information contained in the observable fundamentals. The principal components estimator treats the loadings as being constant over time, that is, $\alpha_{it} \equiv \alpha_i$, and solves the minimisation problem:

$$(\tilde{f}, \tilde{\alpha}_i) = \min_{f,\alpha_i} (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(X_{it} - \alpha_i' f_t\right)^2,$$

where $\tilde{f}$ is a $T \times r$ matrix of common factors and $\tilde{\alpha}$ is an $r \times 1$ vector of factor loadings. By concentrating out α_i and imposing the normalisation constraint $f'f/T = I_r$, the minimisation problem becomes equivalent to maximising $\mathrm{tr}(f'(X'X)f)$, where X is the $T \times N$ matrix of observations. The resulting factor matrix is given by $\sqrt{T}$ times the eigenvectors corresponding to the r largest eigenvalues of the $T \times T$ matrix XX'. It follows from Bates et al.'s (2013) main result that the fundamentals in Equation (9) can be represented through the r principal component estimates, $\tilde{f}_t$, in spite of the structural instability underlying the state vector λ_t. Having obtained the principal component estimates, we estimate the parameters of the state space model (9) by forming the likelihood function:

$$\mathcal{L}_T(\Delta s|\tilde{f};\theta) = -\frac{1}{2}\log(2\pi) - \frac{1}{2T}\log|\Sigma| - \frac{1}{2T}(\Delta s - \mathbb{E}[\Delta s])'\Sigma^{-1}(\Delta s - \mathbb{E}[\Delta s]),$$

where $\Delta s = (\Delta s_1, \cdots, \Delta s_T)'$ with mean $\mathbb{E}[\Delta s] = f\bar{\lambda}$ and variance matrix $\mathbb{V}[\Delta s] = \Sigma$. The parameter vector $\theta = (B(L), \bar{\lambda}, Q, \sigma^2_\epsilon)$ is estimated as

$$\tilde{\theta} = \arg\max_{\theta} \mathcal{L}_T(\Delta s|\tilde{f};\theta).$$

The likelihood can be computed efficiently with the Kalman filter as (9) is a linear state space system. Mikkelsen et al. (2019) show that under standard assumptions and provided $T/N^2 \to 0$, the maximum likelihood estimator is consistent for the parameters of the time-varying factor loadings λ_t, that is, $\tilde{\theta} \xrightarrow{p} \theta$. Once $\tilde{\theta}$ is obtained, the estimates of the factor loadings $\tilde{\lambda}_t$ for $t = 1 \cdots, T$ are computed with the state smoother. As emphasised in Mikkelsen et al. (2019), these estimates are consistent even under missing factors. In addition to the time-varying model, we estimate a constant loadings (CL) benchmark in order to assess the relative contribution of time-varying loadings in explaining exchange rate fluctuations. In that case, equation (7) reduces to $\Delta s_t = f_t \frac{\mu}{\mu - \rho_f}\beta - \frac{1}{\mu - \rho_\kappa}\phi_t$, that is, a present value model for exchange rates with

TABLE 1 Summary statistics exchange rates.

Currency	Mean	Std.Dev.	Min.	Max.	$\hat{\rho}(1)$	$\hat{\rho}(12)$	$\hat{\rho}(24)$	Q_{BP}
AUD	−0.000	0.026	−0.180	0.073	0.342	−0.084	−0.076	0.000
CAD	−0.000	0.017	−0.109	0.062	0.307	−0.091	−0.040	0.000
DKK	0.000	0.023	−0.078	0.062	0.309	−0.093	−0.005	0.000
JPY	0.001	0.025	−0.080	0.103	0.292	−0.035	−0.013	0.000
MXN	−0.005	0.033	−0.321	0.088	0.265	−0.085	0.005	0.000
NZD	0.001	0.026	−0.106	0.074	0.310	0.002	−0.029	0.000
NOK	−0.001	0.026	−0.131	0.057	0.371	−0.107	0.027	0.000
SEK	−0.001	0.026	−0.109	0.071	0.402	−0.104	−0.047	0.000
CHF	0.001	0.025	−0.112	0.081	0.214	−0.063	0.042	0.001
GBP	−0.000	0.022	−0.110	0.059	0.290	−0.036	−0.096	0.000
BRL	−0.033	0.085	−0.368	0.113	0.851	0.486	0.326	0.000
INR	−0.004	0.020	−0.194	0.061	0.200	−0.081	0.019	0.001
ZAR	−0.005	0.035	−0.190	0.152	0.290	−0.078	−0.053	0.000
EUR	−0.000	0.023	−0.079	0.062	0.319	−0.081	−0.021	0.000

Note: Sample period: 1990:05–2021:09. $\hat{\rho}(m)$ denotes the autocorrelation at month m. Q_{BP} denotes the p-value of the Box-Pierce Q_{BP} test. The currency abbreviations stand for Australian Dollar (AUD), Brazilian Real (BRL), Canadian Dollar (CAD), Danish Krone (DKK), Indian Rupee (INR), Mexican Peso (MXN), New Zealand Dollar (NZD), Norwegian Krone (NOK), South African Rand (ZAR), Swedish Krona (SEK), Swiss Franc (CHF), British Pound (GBP), and Euro (EUR).

constant parameters. Therefore, the reduced form relation between fundamentals and exchange rates, $\frac{\mu}{\mu-\rho_f}$, can simply be estimated via OLS by regressing Δs_t on the factors $\tilde{f}_t$. We denote $\frac{\mu}{\mu-\rho_f}\beta$ by $\tilde{\lambda}_{CL}$. Comparing the in- and out-of-sample fit of the two models determines if, indeed, the TVL model fares better at explaining the relationship between exchange rates and fundamentals.

4 | DATA

4.1 | Exchange rate data

We use monthly averages of the US dollar exchange rate vis-à-vis 14 currencies between 1990:05 and 2021:09. The considered exchange rates are as follows: the Australian Dollar (AUD), the Brazilian Real (BRL), the Canadian Dollar (CAD), the Danish Krone (DKK), the Indian Rupee (INR), the Mexican Peso (MXN), the New Zealand Dollar (NZD), the Norwegian Krone (NOK), the South African Rand (ZAR), the Swedish Krona (SEK), the Swiss Franc (CHF), the British Pound (GBP), and the Euro (EUR). The data are compiled from the OECD database.[5] Table 1 reports summary statistics for the first difference of the 14 log exchange rates. Looking at the mean percentage changes, they are all either zero or very close to zero with standard deviations ranging from 1.7% to 8.5%. In terms of fluctuations, the Brazilian Real displays the largest downward movement with −36.8%, whereas the South African Rand appreciated the most over one month with 15.2%. All currencies are positively autocorrelated at one month and with two exceptions negatively correlated at 12 and 24 months. The Box-Pierce test implies that, across currencies, the first three autocorrelations are all statistically significant.

4.2 | Macroeconomic fundamentals and factors

The factors are extracted from a large set of real-time macroeconomic fundamentals. To this end, we combine two different databases. First, we use McCracken and Ng's (2016) FRED-MD database that contains 128 monthly time series of the US economy, categorised into the following: (1) Output & Income, (2) Labour Market, (3) Housing, (4) Orders & Inventories, (5) Money & Credit, (6) Interest Rates, (7) Prices, and (8) Stock Markets. Following McCracken and Ng (2016), five time-series are removed to balance the panel; in addition, we remove the exchange rates in the dataset to exclude them from the factors. In order to ensure stationarity of all variables, we use the same transformations as McCracken and

[5] Prior to 1999, the exchange rate for the ECU is used in place of the Euro, that is, a weighted average of the Austrian Schilling, Belgian and Luxembourg Francs, Finnish Markka, French Franc, German Mark, Irish Pound, Italian Lira, Netherlands Guilder, Portuguese Escudo, and Spanish Peseta.

Ng (2016) and refer to their paper for a detailed description. We collate all real-time vintage releases of the FRED-MD database, starting in 1999:08 and ending in 2022:02. For each vintage of the database, observations are available up to the month that precedes its release date. We remove all time series whose latest observation lags the release date by two months or more. For instance, all time-series ending before 2019:12 are removed from the 2020:01 vintage. As the time lag with which each variable is released changes over time, the number of macro series in the dataset lies between 100 and 128 for each release.

Second, we compile large vintage datasets for the 14 remaining countries from the OECD statistical database. To increase the number of available variables, we collect data for the three largest Euro area economies, Germany, France, and Italy, instead of an aggregate measure. Specifically, for each country, we compile a maximum of nine macro variables, 32 survey indicators, and 3- as well as 10-year yields on government bonds. We collect real-time vintages of this dataset from the OECD's Revision Analysis Database, starting with the first set of vintages released between 1999:02 and 2022:02. As with the FRED-MD data, we remove all variables whose latest observation predates the publication date by more than one month. In many countries, data are published less timely or consistently than in the United States. For example, suppose industrial production data for Mexico are published with a time-lag of two months in 2000:08; that is, the latest available observation for the respective vintage dataset is 2000:06. Then this particular time series will be dropped from the 2000:08 vintage for Mexico to ensure the panel is balanced over time. The first observation in each vintage is recorded in 1990:04. For instance, suppose India only started publishing retail sales data in 1995:05, then no vintage will include this particular variable for India. Therefore, upon removal, the number of macro series available in each vintage version of the dataset lies between 61 and 100; that is, the datasets are balanced across time but unbalanced across countries. We obtain a total of 272 real-time datasets for all countries, the first of which contains $T = 105$ observations per variable, and the last $T = 382$ observations per variable. In each vintage, we group the variables according to the categories in McCracken and Ng (2016). Consistent with McCracken and Ng (2016), the variables are transformed either by taking first log differences or first differences to ensure stationarity. Tables A1 to A6 in the Supporting Information summarise which variables are available for each country and how they are categorised in terms of McCracken and Ng's (2016) classifications.[6] We merge each real-time dataset with the corresponding vintage of the FRED-MD database, yielding 272 real-time vintages with a total number of variables between $N = 168$ and $N = 209$. We use all of these datasets in our out-of-sample estimation, as they replicate the information set available to investors during each month. Thereby, rather than simply using the first release of each observation, we take into account that investors have knowledge of revisions to published data prior to the current month, which could potentially impact their decisions. Caruso and Coroneo (2022) show that this improves forecasts of financial variables relative to just relying on first releases.

For our in-sample exercise, we use the 2022:02 vintage of the FRED-MD and OECD dataset and remove all variables with missing observations before 2021:09. Thus, the time dimension corresponds to $T = 380$. As several variables are released with a lag of more than 1 month, this increases the number of available macro series to $N = 224$. Not every series is available for all countries; hence, the in-sample dataset is also balanced across time but unbalanced across countries.

In practice, the optimal number of factors describing $X_{it} = \alpha_{it'} f_t + \epsilon_{it}$ is not apparent, and multiple factor selection criteria for models where N and T are large exist. Let $V(k) = (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (X_{it} - \tilde{\alpha}_i^{k'} \tilde{f}_t^k)^2$. While setting the number of factors in the model, k, equal to N minimises $V(k)$, it does not imply N corresponds to the optimal number of factors r. Bai and Ng (2002) propose information criteria with a penalty function $g(N, T)$ such that $r = \arg\min_{0 \le k \le k_{max}} \mathrm{IC}(k) = \arg\min_{0 \le k \le k_{max}} \log(V(k)) + kg(N, T)$, where $k_{max} \in \mathbb{N}$ is a maximum number of factors chosen by the researcher (here: $k_{max} = 9$). Due to the penalty term, $r \ll N$. Choi and Jeong (2019) compare the performance of different approaches and suggest using several criteria in combination. We follow their recommendation and first evaluate Bai and Ng's (2002) IC_{p2} and BIC_3, which both pick $r = 9$ factors. Subsequently, we consider several criteria with improved robustness to miss-specification that are found to perform well in Choi and Jeong (2019). Alessi et al. (2010) propose modifications of the penalty functions in Bai and Ng (2002) based around an arbitrary constant, c, as in Hallin and Liška (2007). We set $c \in (0, 10]$, which leads to the conclusion that the optimal number of factors is 1. Kapetanios (2010) suggests a criterion with improved robustness to cross-sectional dependence. When applying the Alessi et al. (2010) modification to this criterion, the optimal number of factors is again found to be 1. In accordance with the advice in Choi and Jeong (2019), we also considered the eigenvalue-based approaches in Ahn and Horenstein (2013). Both the ER and GR test imply $r = 5$,

[6]A complete description of all vintages in the dataset can be found in the Supporting Information.

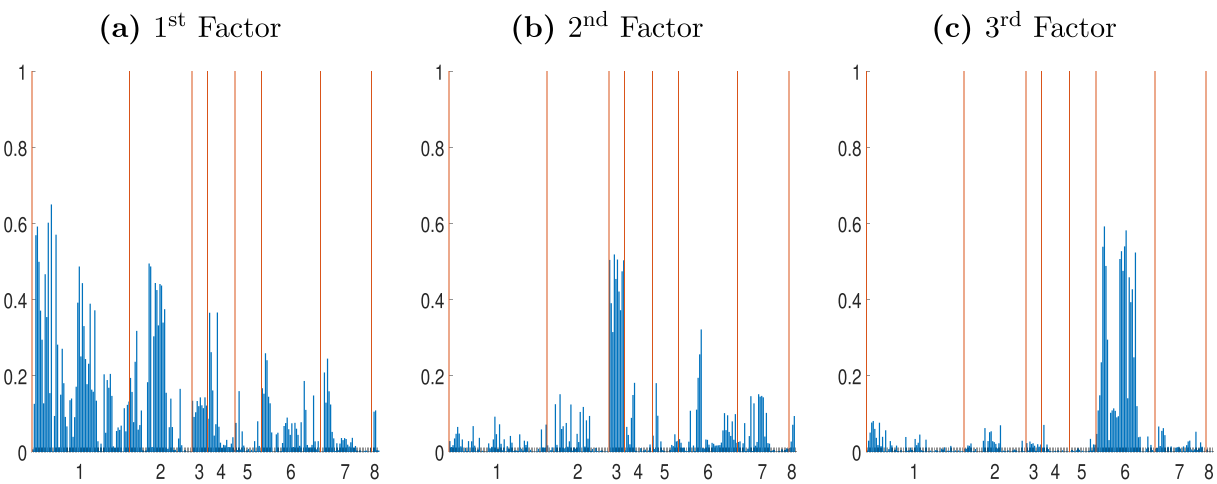

FIGURE 1 Marginal R^2 between factors and macro series. *Note*: R^2 from regression of each series on first, second, and third factors. Series categorised as (1) Output & Income, (2) Labour Market, (3) Housing, (4) Orders & Inventories, (5) Money & Credit, (6) Interest Rates, (7) Prices, and (8) Stock Market.

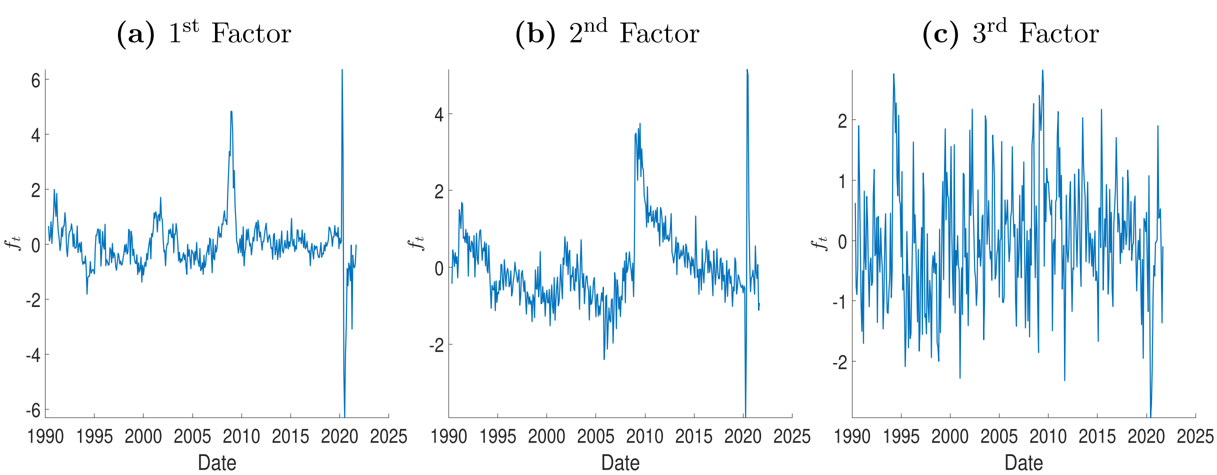

FIGURE 2 Principal components.

suggesting that a low number of factors is indeed a plausible choice. Therefore, one factor and three factors are deemed an appropriate choice for the empirical estimation in this paper.[7]

In Figure 1, we depict the squared correlation of the factors with each macro variable, categorised as described in the figure caption. The displayed correlations are based on the in-sample dataset between 1990:04 and 2021:09. The first factor exhibits strong correlations with measures of output, labour market indicators, and to a lesser extent with manufacturing orders and capacity utilisation. Therefore, we interpret the first factor as an indicator of real economic activity. The second factor correlates mainly with housing data, wherefore we interpret it as a housing factor. Regarding the third factor, it displays strong correlations with interest rates and we interpret it as an interest rate factor.[8]

In Figure 2, the individual factors are plotted. For the first factor, the drop in economic activity during the great recession is clearly visible and so is the COVID-19 crisis towards the end of the sample. The second factor exhibits a structural change after the subprime crisis and large short-dated fluctuations around the COVID-19 crisis. The factor loadings are not sign identified, hence, both the first and second factors are inversely related to the real economy and housing markets. Looking at the third factor, it displays less extreme fluctuations, with an overall minimum in 2020.

[7] Appendix C also reports the results for five factors.

[8] We provide an animation of how the correlations change over time for the vintages used in our out-of-sample exercise in the Supporting Information.

5 | IN-SAMPLE RESULTS

This section presents the empirical results of estimating the state space model in Equation (9) by comparing the constant and the time-varying loadings model in-sample. The discussion focuses on the GBP and the EUR while covering the remaining exchange rates more succinctly. To conduct a comparison of the two models, we use the squared correlation, R^2, between the exchange rate and the in-sample predictions. Furthermore, we report the hit rate (HR) of each model, that is, the percentage of times the model matched the signs of the exchange rate changes. The hit rate indicates how often a model correctly predicts a depreciation or appreciation. The two criteria are shown in Table 2. In addition, the table reports the p-values of a likelihood-ratio test with a null hypothesis of no significant differences in the likelihoods of the two models.

Consider first the estimates for one factor, the real economy factor. The top panel of Figure 3 shows the results for the GBP and the EUR: the common component obtained from the TVL model (blue), the CL model (red), and the actual exchange rate changes (black). Particularly during the great recession, the time-varying model can capture the fluctuations in the exchange rate better. This is reflected in the R^2, according to which the model can explain 35% (23%) of the variation in the EUR (GBP). It assigns accurate directional changes in 73.7% of the cases for the EUR and 60.5% for the GBP. In contrast, the CL model only has an R^2 of 1% and 3%, respectively, i.e. it has almost no explanatory power. With 47.5%, the hit rate of the CL model for the EUR is as good as random, the same holds for the GBP with 52.3%. It should be noted that the CL results for the GBP are among the highest out of the 14 exchange rates. The lowest are the ones for CHF and JPY where 0% of the fluctuations are explained. Looking at the time-varying model, it has the highest explanatory power for the MXN with an R^2 of 54%, and the lowest R^2 for the INR (1%). In case of the latter, the TVL model and the CL model have the same R^2 which suggests the real economy factor offers no explanatory power for the INR. Nevertheless, the time-varying one-factor model adds substantial explanatory power over the model with constant coefficients for all other currencies, as it consistently outperforms the CL model according to the two metrics. This conclusion is supported by the likelihood ratio tests. With three exceptions in case of the one-factor model (INR, CHF, and JPY), the test always finds that the likelihood of the CL model is significantly smaller. In a next step, we also include the interest rate and the housing factor into the model. The actual and fitted values for the GBP and the EUR are presented in the bottom panel of Figure 3. In particular for the EUR, the fit of the time-varying model improves visibly—the model tracks the

TABLE 2 In-sample performance.

	1 factor					3 factors				
	R^2		Hit rate		LR	R^2		Hit rate		LR
Currency	TVL	CL	TVL	CL	p-val.	TVL	CL	TVL	CL	p-val.
AUD	0.44	0.01	74.54	49.60	0.00	0.53	0.04	75.86	56.76	0.00
CAD	0.32	0.02	70.29	52.52	0.00	0.36	0.04	68.70	53.32	0.00
DKK	0.33	0.01	70.29	47.21	0.00	0.61	0.01	79.58	54.91	0.00
JPY	0.09	0.00	57.03	52.25	0.19	0.41	0.00	67.90	53.58	0.01
MXN	0.54	0.01	71.35	54.38	0.00	0.55	0.02	68.44	50.40	0.00
NZD	0.27	0.02	70.03	49.34	0.00	0.51	0.05	78.78	57.56	0.00
NOK	0.28	0.02	74.01	46.95	0.00	0.41	0.03	75.07	52.79	0.00
SEK	0.23	0.03	74.27	50.66	0.00	0.48	0.04	76.66	54.91	0.00
CHF	0.37	0.00	69.50	44.83	0.07	0.64	0.01	82.76	52.79	0.06
GBP	0.23	0.03	60.48	52.25	0.00	0.27	0.04	65.52	57.29	0.01
BRL	0.50	0.01	72.94	53.58	0.00	0.86	0.01	86.47	55.70	0.00
INR	0.01	0.01	50.13	50.13	1.00	0.88	0.01	89.39	52.25	0.00
ZAR	0.39	0.03	69.76	50.40	0.00	0.48	0.06	76.13	55.44	0.00
EUR	0.35	0.01	73.74	47.48	0.00	0.61	0.01	81.70	53.85	0.00
Mean	0.31	0.01	68.45	50.11		0.54	0.03	76.64	54.40	
Median	0.32	0.01	70.29	50.27		0.52	0.02	76.39	54.38	

Note: The table reports measures of in-sample fit to compare the TVL model in Equation (9) and the CL model. Namely, both the squared correlations between changes in the exchange rate and the in-sample prediction of the TVL & CL model as well as the hit rate in %. The latter is the fraction of times the sign of the fitted values corresponded to the sign of the realised values. In addition, the table reports the p-values of a likelihood ratio test (LR) between the two models. The currency abbreviations stand for Australian Dollar (AUD), Brazilian Real (BRL), Canadian Dollar (CAD), Danish Krone (DKK), Indian Rupee (INR), Mexican Peso (MXN), New Zealand Dollar (NZD), Norwegian Krone (NOK), South African Rand (ZAR), Swedish Krona (SEK), Swiss Franc (CHF), British Pound (GBP), and Euro (EUR).

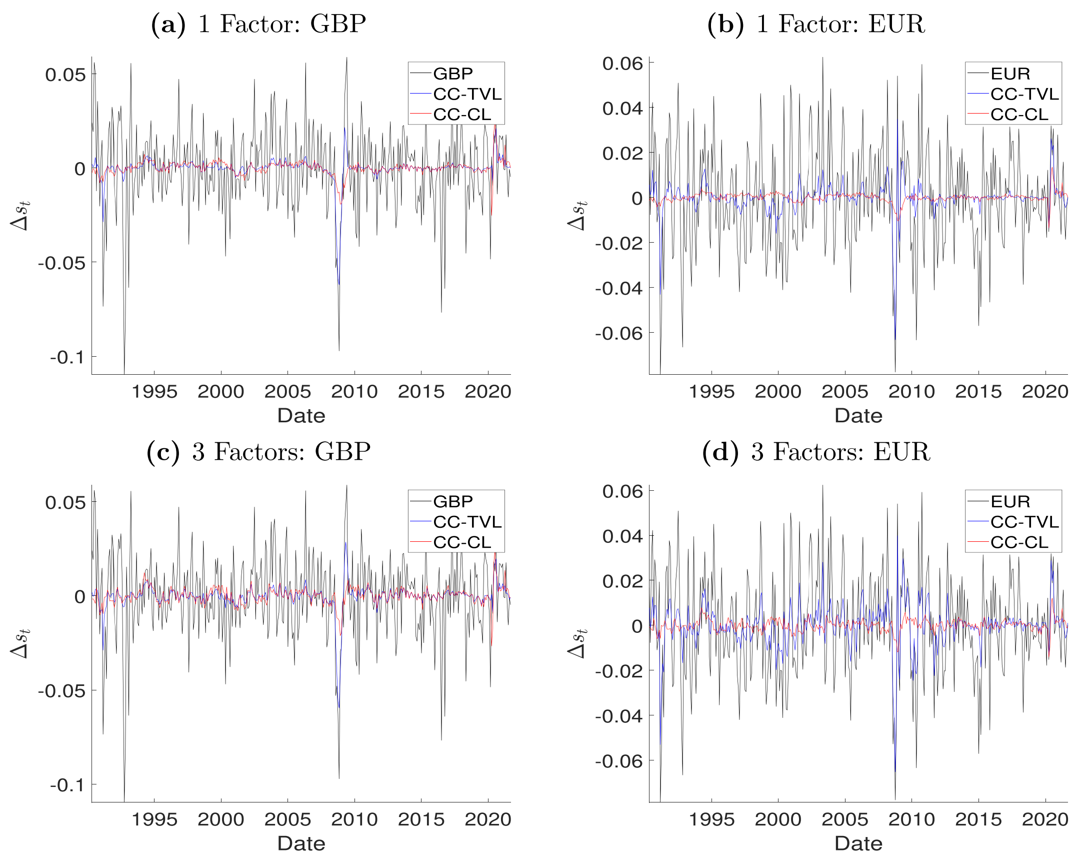

FIGURE 3 In-sample fit. *Note*: The figure displays the in-sample fit for the Euro (EUR) and the British Pound (GBP) of a model with one factor and three factors. The black line is the FX change, the blue line is the TVL model fit, and the red line is the CL model fit. CC-TVL stands for the Common Component of the State Space model, CC-CL for the Common Component of the CL model.

depreciation during the Euro-crises in 2010, 2012, and 2015 remarkably well. The same holds true for the early 2000s. The fit for the GBP also appears to have improved, even though to a lesser extend. Note that, as before, the GBP exhibits the worst fit of all time-varying regressions with an R^2 of 27% followed by the CAD with 36%. Still, it manages to predict whether the exchange rate appreciates or depreciates in 65.5% of all cases. Regarding the EUR, the explanatory power of the time-varying model has jumped to an R^2 of 61%, and the hit rate corresponds to 81%. On the other hand, the CL model still only manages to explain 1% of the variations in the EUR and 4% in the GBP. Overall, the time-varying model of the INR has now the highest R^2 and hit rate with 88% and 89.4%, respectively—having displayed the worst fit in the one-factor model. This implies that only the second and third factors offer predictive ability for this currency. The INR is followed by BRL, CHF, EUR, and DKK for which the explanatory power always exceeds 60%. Generally, we see an improvement in the R^2 across exchange rates—the hit rate declines slightly for the MXN in the time-varying framework but is close to 80% in most cases; however, looking at the CL fit, it is still only slightly better than 50% and never above 58%. After the INR, the BRL exhibits the greatest improvement with the R^2 increasing to 86% from 50%. We graph the differences between the one- and three-factor models in Figure 4. Both INR and BRL were hit by a currency crisis at the beginning of the sample period. Regardless of whether one factor or three factors are used, the CL model can map neither fluctuations in INR nor BRL. In contrast, the TVL model with three factors is able to capture both currency crises as well as subsequent variations. For the remaining series, the one-factor model is also able to capture exchange rate fluctuations better around the financial crisis (see Appendix B). As is the case with the EUR, GBP, INR, and BRL, the three-factor model substantiates the ability of the time-varying model to outperform the constant loadings framework.

5.1 | Instabilities in factor loadings

This subsection considers the role of parameter instability in greater detail. First, we revisit the GBP and the EUR and discuss the time-varying loadings on the real economy factor, depicted in the left panel of Figure 5.

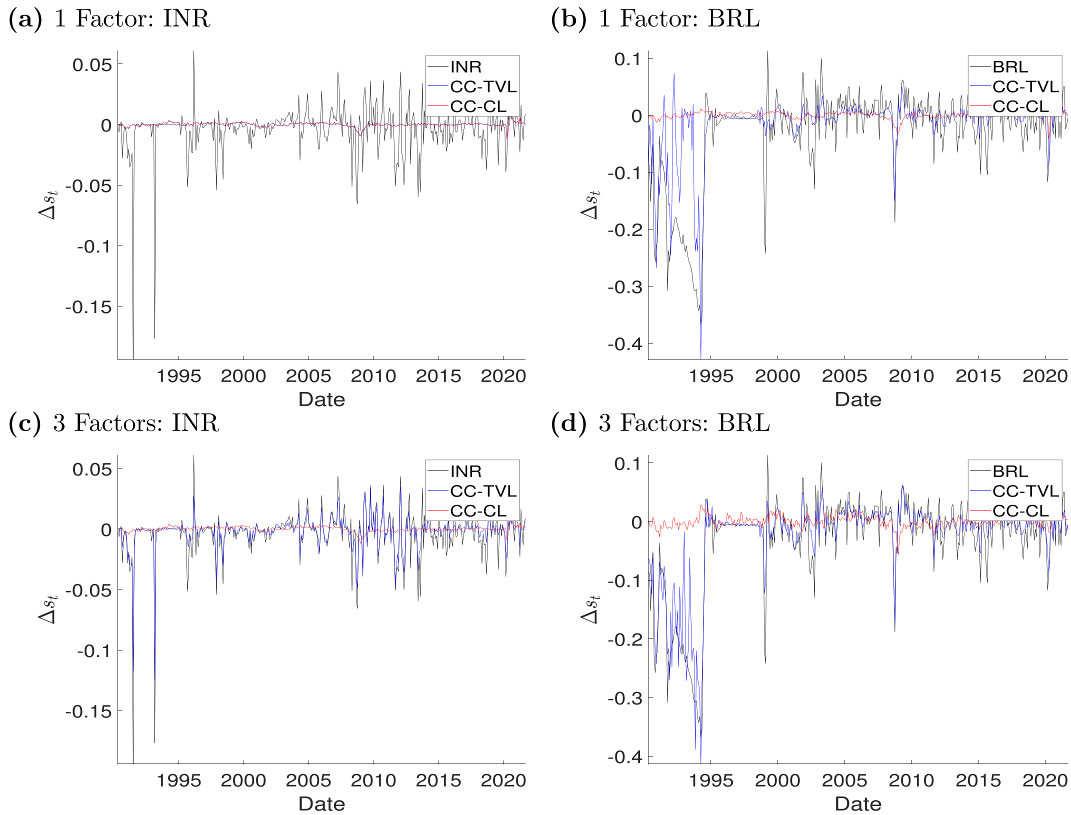

FIGURE 4 In-sample fit—BRL and INR. *Note:* The figure displays the results for the Brazilian Real (BRL) and the Indian Rupee (INR) of a model estimated with one factor and three factors. The black line is the FX change, the blue line is the TVL model fit, and the red line is the CL model fit. CC-TVL stands for the Common Component of the State Space model, CC-CL for the Common Component of the CL model.

The dashed red lines correspond to the confidence intervals of the constant loadings estimates. For both currencies, we observe a high degree of variation in the time-varying loadings, especially in 2008–2009. In the GBP model, the loadings decline first and then rise sharply; therefore, they are considerably outside the CL confidence interval during the financial crisis. This is a consistent theme across exchange rates.[9] The loading in the EUR model crosses the CL confidence bands more often, displaying large fluctuations throughout the sample period. However, the EUR CL estimate itself is insignificant. Contrary to the GBP, the loadings exhibit large positive spikes in the early 2000s, consistent with the EUR depreciation *vis-à-vis* the dollar during that period. Although the loadings are not sign identified, Figure 5 shows that the loadings display frequent sign changes, in the sense that they fluctuate significantly above and below the OLS confidence bands. This points to significant instabilities in the factor loadings that the CL model cannot capture. The right panel of Figure 5 plots the loadings on the first three factors. It can be seen that the explanatory power of the second factor differs across the two currencies with the magnitude of the loadings being much greater for the EUR. What is more, as the methodology in Mikkelsen et al. (2019) is robust to missing factors, the first loadings in the three-factor model are always identical to the loadings in the one-factor model.

5.2 | Instabilities in factor structures

Changes in factor structures occur if either (i) the number of factors changes, (ii) related to this, some factors disappear and others appear, or (ii) the direction of the influence of a factor changes. Time-varying loadings capture a reduction in the number of factors, since loadings can be (close to) zero for periods. They also capture changes in the direction of influence, since the sign of the loadings can change. The right panel of Figure 5 shows that in case of the GBP, the second loadings are always close to zero apart from few deviations. For both currencies, the third loadings are constant and only marginally above zero, meaning only the first and second factors affect these currencies. Next, we look at three

[9]The plots of the first factor loadings for the remaining 12 currencies are reported in Appendix B.

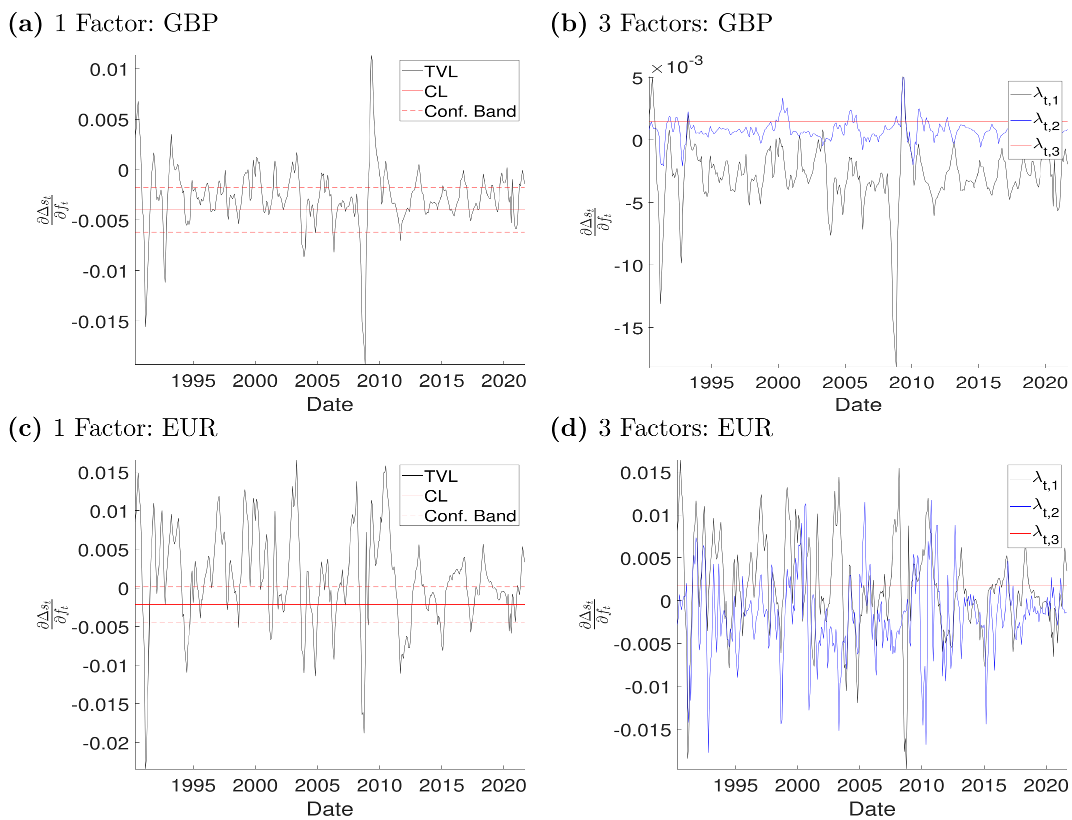

FIGURE 5 Loadings—GBP & EUR. *Note*: The left panel of the figure displays the loadings on the real economy factor. The black line is the time-varying loading (TVL), the red line the constant loadings (CL) estimate, and the dashed lines are the CL confidence intervals. The right panel displays the loadings of the three-factor model. The black line is the loadings on the real economy factor, the blue line the loadings on the housing factor, and the red line the loadings on the interest rate factor.

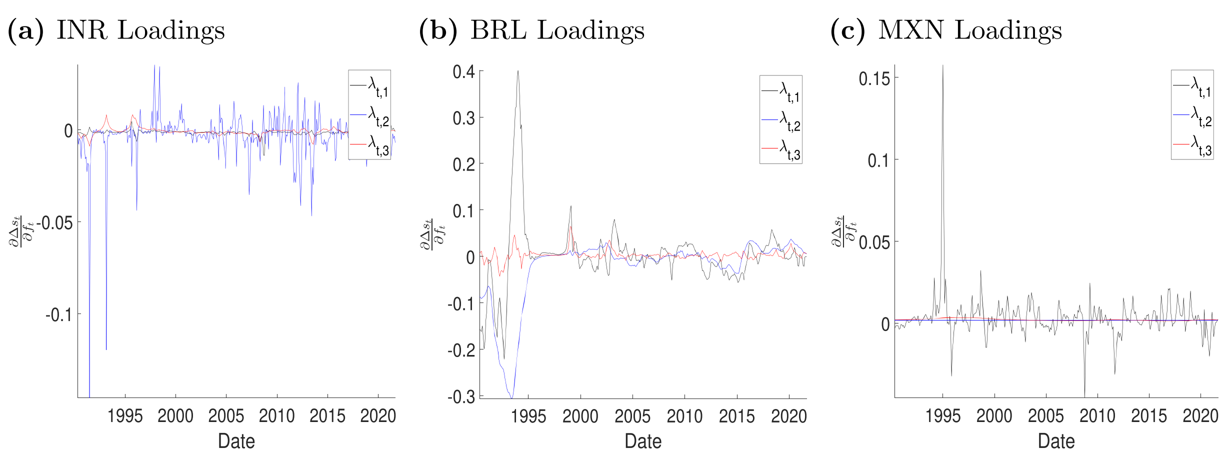

FIGURE 6 Loadings—unstable factor structure. *Note*: The figure displays the loadings of the three-factor model. The black line are the loadings on the real economy factor, the blue line the loadings on the housing factor, and the red line the loadings on the interest rate factor.

emerging market currencies that were subject to crises at the beginning of the sample period. Therefore, they display greater instabilities in both loadings and factor structure. Figure 6 plots the loadings on the first three factors for INR, BRL, and MXN. Several insights emerge from the figure. First, the dominant loadings differ from currency to currency. While the first loadings are virtually zero for the INR, the second loadings exhibit large swings and the third loadings show only temporal deviations from zero. With respect to the latter, the same is true for the BRL. However, both the second and third factors load heavily onto the BRL during the early 1990s, while fluctuating less drastically in the 2000s where the loadings are close to zero in some months. This contrasts with the MXN, for which only the first loadings are markedly different from zero, with large spikes during the Peso crisis and the financial crisis.

Unlike the CL model, the TVL framework can account for such changes in the importance of certain factors. It is therefore able to model instabilities in both loadings and factor structures which explains the superior fit discussed above.

6 | OUT-OF-SAMPLE RESULTS

This section compares the performance of the two models, CL and TVL out-of-sample using the three-factor model. First, we elaborate on the chosen forecast evaluation methods; and second, assess the forecasting results. Specifically, we compare (i) the relative predictive ability of the two models, (ii) their direction accuracy, that is, how well they forecast an appreciation or depreciation, and (iii) their relative performance over time. For the out-of-sample estimation, we use the large vintage datasets described above.

6.1 | Forecast setup and evaluation

To forecast Δs_{t+h}, where h is the forecast horizon, we divide the sample into in-sample and out-of-sample portions. We denote the total number of observations by T, the number of in-sample observations by R, and the number of out-of-sample predictions by P, so $T = R + P + h + 1$. We generate direct h-step ahead forecasts according to the following specification[10]:

$$\Delta s_{t+h|t} = f'_{t-1+h|t} \lambda_{t+h|t} + e_{t+h|t}$$

where $t = R, \cdots, T-h$ and $e_{t+h|t}$ is the forecast error. Note that although the factors, $f_{t-1+h|t}$, are extracted from data one date prior to the observed returns, the estimates are nonetheless conditional on the time-t information set. This is due to the fact that data are released with a lag, as, for example, inflation figures for March are published in April. Our setup ensures the model is truly real-time by accounting for the fact that exchange rates can react to data only upon release. As our dataset includes over 20 survey indicators, we also incorporate the impact of expectations that anticipate some of this response. We set $h = \{1, 6, 12\}$ and use a rolling window to compute P predictions. Forecasts for the loadings, $\hat{\lambda}_{t+h|t}$, are easily obtained as the one-step-ahead predictions of the Kalman filter. The Kalman filter generates optimal forecasts as it, by construction, minimises the mean squared error (MSE) between predictions and observations. For the constant loadings benchmark, λ does not need to be forecast as it simply corresponds to the CL estimates at each iteration. Regarding the factors, we recompute the principal components for each of the 272 real-time datasets. To generate h-step ahead forecasts of the factors, we fit a VAR(1) to the real-time estimates at each of the P forecasting steps and use the coefficients to forecast $\hat{f}_{t-1+h|t}$. One then obtains the out-of-sample estimates as $\widehat{\Delta s}_{t+h|t} = \hat{f}'_{t-1+h|t} \hat{\lambda}_{t+h|t}$ and $\widetilde{\Delta s}_{t+h|t} = \hat{f}'_{t-1+h|t} \hat{\lambda}_{CL}$. Both forecasts are only based on the actual information set available in real-time at each period. To assess whether time-varying loadings also improve the out-of-sample fit of the factor model, we compare the TVL and CL forecasts using several tests. In addition we compare both models to a random walk (RW) for the level of the log exchange rate, that is, $\mathbb{E}[s_{t+h|t} - s_t] = 0$.

Selecting adequate tests of predictive ability is of paramount importance in out-of-sample evaluation. The properties of tests for nested models are different because their forecast errors converge asymptotically (Clark & McCracken, 2001). Furthermore, window choice is an important determinant in forecast evaluation. A large P provides more forecast information, while a large R improves parameter accuracy. In fact, Mikkelsen et al. (2019) show through Monte-Carlo simulations that in order for the bias in the autoregressive parameters of the loadings to be below 10%, the estimation sample should be $R \geq 200$. However, a litmus test for every exchange rate forecast is the financial crisis. To put the model to this test, the forecasts need to be evaluated using a criterion that is robust to in-sample estimation errors, as one would have $R < 200$ for a prediction window starting prior to the crisis. Giacomini and White (2006) propose a test of Conditional Predictive Ability (CPA) that introduces estimation error under the null hypothesis. Define the forecast loss differential between TVL and CL as $\left\{\Delta\mathcal{L}_t(\hat{f}'_{t-1}\hat{\lambda}_t, \hat{f}'_{t-1}\hat{\lambda}_{CL})\right\}_{t=R+h}^{T} = \left\{\mathcal{L}(\Delta s_t, \hat{f}'_{t-1}\hat{\lambda}_t) - \mathcal{L}(\Delta s_t, \hat{f}'_{t-1}\hat{\lambda}_{CL})\right\}_{t=R+h}^{T}$, where $\mathcal{L}(\cdot)$ is a forecast loss function. The null hypothesis of the CPA test is

$$\mathcal{H}_0 : \mathbb{E}\left[\Delta\mathcal{L}_t(\hat{f}'_{t-1}\hat{\lambda}_t, \hat{f}'_{t-1}\hat{\lambda}_{CL})|\mathcal{F}_t\right] = 0$$

[10] See Boivin and Ng (2005) for a comparison of different approaches to generating factor-based forecasts.

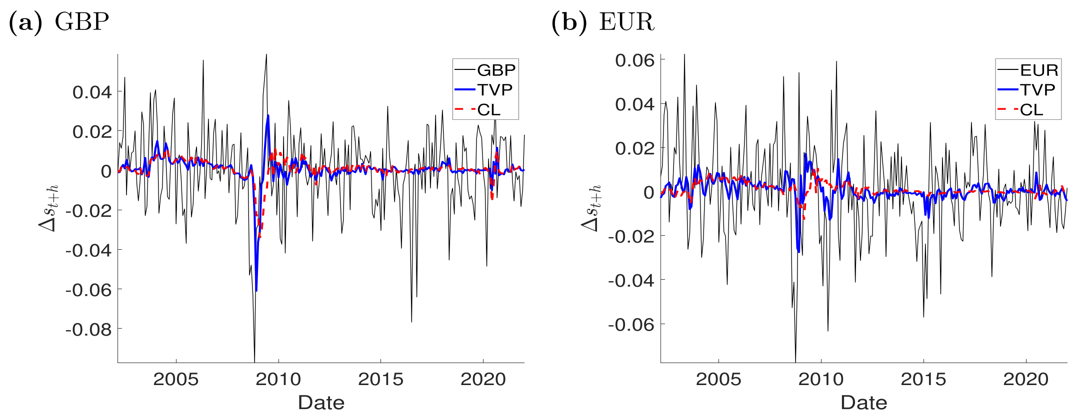

FIGURE 7 Rolling window forecast, $h = 1$. *Note:* This figure plots the out-of-sample, one-step-ahead, rolling window forecasting results of the time-varying loadings (TVL, blue line) and the constant loadings (CL, red line) model. The black line corresponds to the actual exchange rate.

and $\mathcal{F}_t$ is the time-t information set available to the forecaster. For the implementation of the test, we condition on lagged values of the loss differential. In addition, we also use Giacomini and White's (2006) unconditional predictive ability (UPA) test.[11] As both tests are valid for nested as well as for non-nested models and their asymptotic properties are derived for $R < P \to \infty$, they are well-suited in this application. We choose $P = 238$ and $R = 142$ as the baseline configuration and report additional forecasts in Appendix C.

Leitch and Tanner (1991) argue that, while one model may produce a smaller forecasting error than another model, it can still perform worse when it comes to predicting sign changes. In the case of exchange rates, a desirable feature of a model is its ability to forecast an appreciation or depreciation. To assess this statistically, we use Pesaran and Timmermann's (1992) nonparametric test of predictive performance. The test compares the signs of the predicted and realised values and, in doing so, uses no additional information. Thus, it does not require knowledge of the underlying probability distribution of the forecast. Although the test does not put two models in relation to one another, it indicates which model is able to identify a higher number of predictable relationships.

Given the structural instabilities in the exchange rate regression, it may well be the case that the relative forecasting performance of the models is itself unstable. Indeed, Rossi (2013) finds that the forecasting power of many exchange rate models breaks down over time. Notably, however, parameter instability itself does not necessarily engender unstable relative forecast performance.[12] While the Giacomini and White (2006) test selects the best global model, Giacomini and Rossi (2010) propose a fluctuation test that compares the performance of two competing models at each point in time and allows for nested models by adopting the same asymptotic framework as Giacomini and White (2006). The test statistic is computed over a rolling window and equal to

$$GR_{t,m} = \hat{\sigma}^{-1} m^{-1/2} \sum_{j=t-m/2}^{t+m/2-1} \Delta\mathcal{L}_j(\hat{f}'_{j-1,R}\hat{\lambda}_{j,R}, \hat{f}'_{j-1,R}\hat{\lambda}_{CL})$$

where $t = R + h + m/2, \cdots, T - m/2 + 1$, $\hat{\sigma}$ is the HAC estimator of the variance of the loss differential, and m is the size of the rolling window over which it is computed.

6.2 | Forecast results

Focusing foremost on GBP and EUR, we now discuss the forecasting results. Figure 7 depicts the TVL and CL forecasts for the two currencies as well as the realised values. At first glance, it appears the time-varying model performs slightly better for the GBP during the financial crisis—and for the EUR also during the subsequent years. Figure 8 plots the forecasts of the remaining series which paint a similar picture; in particular, the INR (Figure 8f) is forecast remarkably well.

[11]The UPA test is identical to the popular Diebold and Mariano (1995) test, but derived under different assumptions that render it valid for our purposes.

[12]For a detailed discussion of forecasting under instabilities, see Rossi (2021).

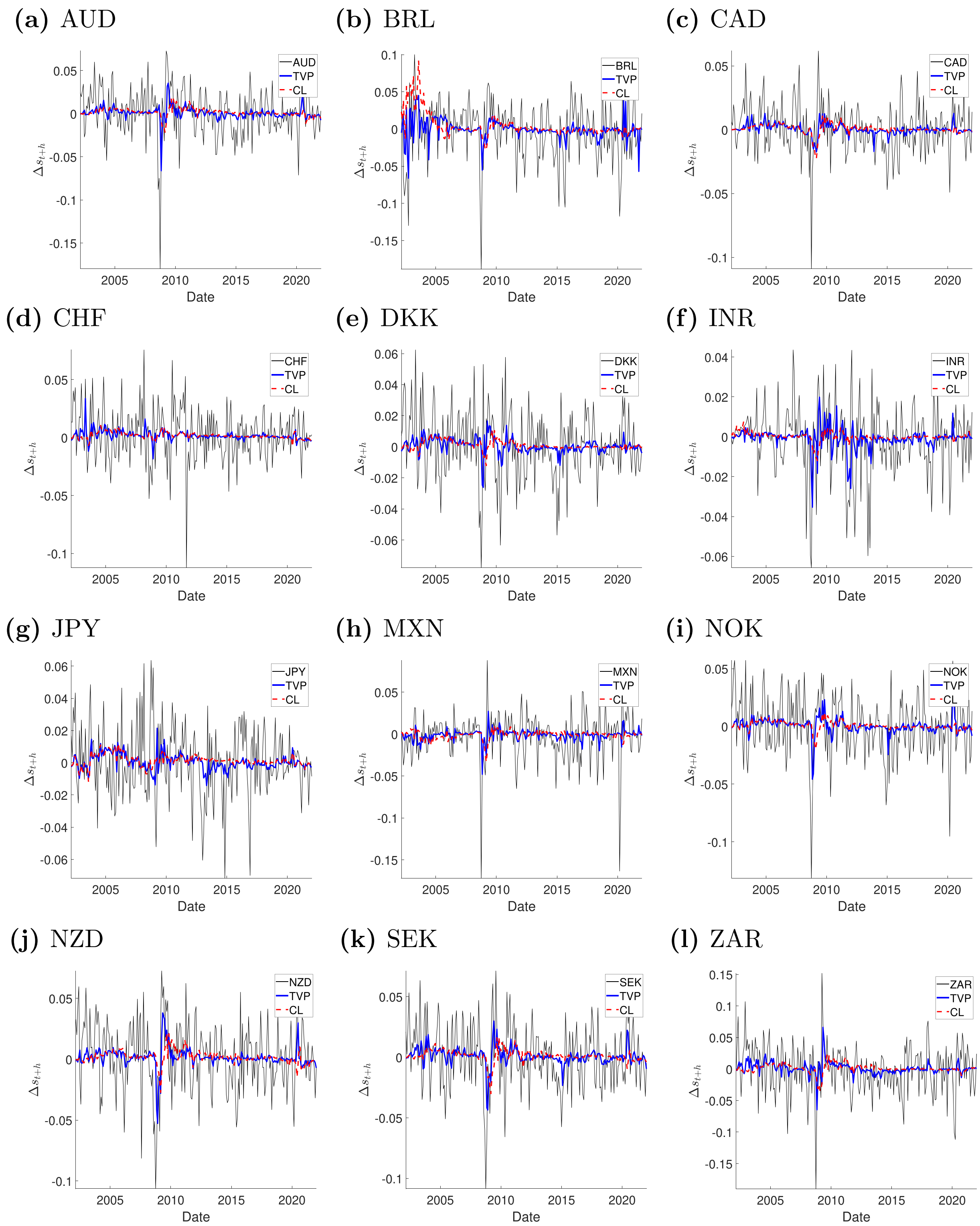

FIGURE 8 Rolling window forecast, $h = 1$. *Note:* This figure plots the out-of-sample, one-step-ahead, rolling window forecasting results of the time-varying loadings (TVL, blue line) and the constant loadings (CL, red line) model. The black line corresponds to the actual exchange rate.

We report statistical forecast accuracy tests in Table 3. Panel A presents the results for one-step-ahead predictions. Columns 2 and 3 of the table report the root mean square error (RMSE) of the forecasts relative to the RMSE of a random walk forecast. That is, if the value in columns 2 and 3 is smaller than one, the respective model has a lower RMSE smaller than a random walk. The TVL model achieves an RMSE that is between 4.4% and 0.3% smaller than a random walk for 8

TABLE 3 Forecast statistics.

	RMSE		TVL vs. CL						TVL vs. RW				CL vs. RW			
			Quad.		Abs.		DA		Quad.		Abs.		Quad.		Abs.	
	TVL	CL	CPA	UPA	CPA	UPA	TVL	CL	CPA	UPA	CPA	UPA	CPA	UPA	CPA	UPA
Panel A: 3 factors, $h = 1$																
AUD	0.972	1.012	0.075	0.024	0.117	0.086	0.029	0.717	0.252	0.119	0.237	0.104	0.502	0.281	0.717	0.931
CAD	0.985	1.004	0.177	0.065	0.329	0.136	0.017	0.187	0.133	0.225	0.340	0.226	0.545	0.752	0.991	0.907
DKK	1.002	1.015	0.445	0.431	0.238	0.110	0.020	0.685	0.698	0.940	0.380	0.566	0.148	0.114	0.240	0.176
JPY	1.027	1.029	0.507	0.908	0.849	0.819	0.323	0.712	0.168	0.088	0.374	0.151	0.042	0.011	0.262	0.096
MXN	1.004	1.019	0.411	0.398	0.053	0.051	0.006	0.429	0.169	0.775	0.203	0.285	0.413	0.202	0.478	0.290
NZD	0.966	1.025	0.029	0.011	0.049	0.055	0.036	0.378	0.217	0.123	0.278	0.173	0.234	0.108	0.369	0.494
NOK	0.976	1.009	0.047	0.033	0.037	0.029	0.003	0.680	0.266	0.166	0.545	0.343	0.090	0.220	0.146	0.074
SEK	0.956	1.014	0.005	0.009	0.003	0.010	0.001	0.099	0.008	0.034	0.023	0.082	0.262	0.283	0.293	0.252
CHF	1.004	1.009	0.673	0.691	0.746	0.949	0.914	0.830	0.938	0.775	0.676	0.409	0.022	0.276	0.180	0.209
GBP	0.985	1.023	0.258	0.136	0.538	0.511	0.178	0.454	0.722	0.524	0.358	0.597	0.149	0.232	0.262	0.247
BRL	0.972	1.090	0.004	0.012	0.004	0.006	0.123	0.509	0.408	0.331	0.212	0.373	0.047	0.021	0.092	0.041
INR	0.993	1.002	0.447	0.766	0.919	0.811	0.211	0.039	0.429	0.823	0.960	0.950	0.983	0.851	0.878	0.663
ZAR	0.997	1.023	0.099	0.039	0.202	0.078	0.113	0.493	0.615	0.885	0.862	0.581	0.196	0.098	0.171	0.128
EUR	1.000	1.015	0.349	0.420	0.252	0.116	0.020	0.751	0.557	0.994	0.353	0.567	0.186	0.131	0.265	0.176
Σ	9	0	4	6	4	3	8	1	1	1	1	0	3	2	0	1
Panel B: 3 factors, $h = 6$																
AUD	0.996	1.013	0.254	0.209	0.611	0.328	0.102	0.901	0.027	0.195	0.087	0.035	0.084	0.102	0.073	0.631
CAD	1.017	1.020	0.179	0.729	0.422	0.736	0.076	0.565	0.101	0.195	0.163	0.324	0.097	0.098	0.146	0.105
DKK	1.002	1.007	0.299	0.503	0.732	0.513	0.399	0.487	0.057	0.677	0.225	0.616	0.030	0.178	0.142	0.090
JPY	1.015	1.012	0.713	0.482	0.967	0.801	0.444	0.668	0.040	0.015	0.425	0.215	0.059	0.032	0.243	0.111
MXN	1.003	1.011	0.581	0.317	0.413	0.411	0.712	0.994	0.784	0.563	0.105	0.286	0.103	0.152	0.034	0.089
NZD	1.004	1.019	0.435	0.342	0.730	0.468	0.980	0.972	0.293	0.517	0.771	0.844	0.127	0.066	0.129	0.219
NOK	0.996	1.007	0.276	0.129	0.270	0.114	0.120	0.596	0.015	0.311	0.151	0.160	0.095	0.247	0.122	0.271
SEK	1.003	1.025	0.485	0.261	0.055	0.115	0.071	0.758	0.062	0.530	0.190	0.690	0.262	0.140	0.092	0.035
CHF	1.009	1.000	0.094	0.056	0.114	0.107	0.448	0.382	0.211	0.124	0.448	0.355	0.299	0.972	0.637	0.905
GBP	1.008	1.022	0.321	0.132	0.533	0.304	0.551	0.680	0.373	0.364	0.382	0.192	0.195	0.078	0.257	0.105
BRL	1.006	1.035	0.359	0.092	0.540	0.282	0.000	0.049	0.850	0.669	0.276	0.562	0.112	0.116	0.075	0.433
INR	0.993	1.003	0.395	0.189	0.792	0.727	0.273	0.026	0.215	0.239	0.557	0.503	0.375	0.452	0.892	0.881
ZAR	1.002	1.020	0.342	0.127	0.196	0.060	0.487	0.980	0.732	0.647	0.966	0.946	0.031	0.011	0.029	0.008
EUR	1.003	1.006	0.359	0.611	0.595	0.572	0.598	0.597	0.059	0.621	0.166	0.616	0.034	0.221	0.158	0.124

Continues

exchange rates and between 0.2% and 2.7% greater for four exchange rates. In contrast, the RMSE of the CL model is always between 0.2% and 2.9% larger. We proceed by comparing the TVL and CL models directly. Columns 4 and 5 contain the p-values of the CPA and UPA test, computed for a quadratic loss differential. At the 5% level, the former rejects four times in favour of the TVL model, and the latter six times. We conduct the same tests using a quadratic loss differential (Columns 6 and 7) for which they reject four and three times, respectively, always in favour of the TVL model. In Columns 8 and 9, we report the p-values of the direction accuracy test, which rejects the null hypothesis of no directional accuracy 8 times for the TVL model and only once for the CL model. All three tests provide consistent evidence of improved predictive ability when time variation is accounted for. Next, we compare the TVL model to random walk forecasts using the CPA and UPA tests (p-values reported in Columns 10 to 13). We do not report the direction accuracy test, as the random walk forecast, by definition, has zero directional accuracy. With a quadratic loss function, both CPA and UPA reject once in favour of the TVL model and never in favour of the random walk. For an absolute loss function, the CPA test rejects once and again in favour of the TVL model. Finally, we compare the CL model with the random walk (p-values reported in Columns 14 to 17). Now, the CPA test rejects three times, and the UPA test twice, in favour of the random walk. Using an absolute loss function, the UPA test rejects once. The results imply that the TVL model does not only have better predictive ability than the CL model but also that it produces better forecasts than the CL model when comparing both to a random walk. We repeat this exercise for a forecast horizon of $h = 6$ in panel B of Table 3 and for $h = 12$ in panel C. While TVL and CL have equal predictive ability for $h = 6$ according to the CPA and UPA test, the CPA test based on a quadratic loss function now rejects the null hypothesis that the TVL model and a random walk have equal predictive ability three times: twice in favour of the TVL model, for AUD and NOK, and in case of the JPY in favour of the random walk. The UPA test also

TABLE 3 Continued.

	RMSE		TVL vs. CL						TVL vs. RW				CL vs. RW			
			Quad.		Abs.		DA		Quad.		Abs.		Quad.		Abs.	
	TVL	CL	CPA	UPA	CPA	UPA	TVL	CL	CPA	UPA	CPA	UPA	CPA	UPA	CPA	UPA
∑	3	0	0	0	0	0	1	2	3	1	0	1	3	2	2	2
Panel C: 3 factors, $h = 12$																
AUD	0.999	1.004	0.274	0.172	0.388	0.428	0.124	0.992	0.044	0.705	0.338	0.145	0.655	0.380	0.445	0.894
CAD	1.035	1.005	0.333	0.188	0.218	0.113	0.142	0.425	0.181	0.168	0.313	0.131	0.532	0.416	0.728	0.699
DKK	1.009	0.997	0.215	0.209	0.507	0.189	0.362	0.333	0.052	0.141	0.058	0.237	0.117	0.538	0.336	0.458
JPY	1.006	1.003	0.464	0.170	0.155	0.150	0.916	0.333	0.387	0.206	0.054	0.051	0.598	0.552	0.336	0.157
MXN	1.022	1.004	0.527	0.322	0.478	0.367	0.214	0.998	0.140	0.179	0.077	0.031	0.159	0.215	0.035	0.009
NZD	1.004	1.001	0.407	0.441	0.588	0.841	0.987	0.978	0.079	0.246	0.948	0.745	0.951	0.923	0.556	0.736
NOK	0.998	0.999	0.046	0.840	0.269	0.549	0.099	0.525	0.009	0.716	0.052	0.696	0.140	0.814	0.562	0.962
SEK	1.002	1.002	0.971	0.953	0.258	0.120	0.075	0.236	0.024	0.614	0.019	0.687	0.148	0.684	0.423	0.399
CHF	1.020	0.998	0.215	0.243	0.258	0.135	0.555	0.367	0.417	0.216	0.456	0.206	0.403	0.660	0.436	0.416
GBP	1.002	0.996	0.127	0.119	0.092	0.147	0.710	0.784	0.435	0.812	0.642	0.387	0.878	0.671	0.886	0.994
BRL	0.993	0.997	0.168	0.625	0.687	0.978	0.006	0.154	0.651	0.422	0.271	0.269	0.563	0.775	0.489	0.371
INR	0.995	1.000	0.426	0.317	0.829	0.531	0.915	0.348	0.343	0.415	0.884	0.977	0.948	0.940	0.501	0.523
ZAR	0.998	1.007	0.065	0.012	0.467	0.155	0.920	1.000	0.241	0.582	0.773	0.809	0.062	0.018	0.080	0.192
EUR	1.009	0.996	0.314	0.188	0.385	0.134	0.621	0.433	0.045	0.134	0.061	0.188	0.106	0.467	0.302	0.397
∑	5	6	1	1	0	0	1	0	4	0	1	1	0	1	1	1

Note: Columns 2 and 3 report the root mean square error (RMSE) of TVL and CL forecasts divided by the RMSE of forecasts by a random walk (RW). A value <1 implies the respective model has a smaller RMSE than the RW. Columns 4 to 9 compare the TVL and CL models and report the *p*-values of the conditional predictive ability (CPA) and unconditional predictive ability (UPA) test of Giacomini and White (2006) using a quadratic (Quad.) and an absolute (Abs.) loss function. Further, they report the *p*-values of the Pesaran and Timmermann (1992) nonparametric Direction Accuracy (DA) test for TVL and CL. Columns 10 to 13 compare TVL and RW, reporting *p*-values of the CPA and UPA test for quadratic and absolute loss differentials. Columns 14 to 17 compare CL and RW, reporting *p*-values of the same tests. The rows denoted by ∑ report the total number of *p*-values ≤ 5%, and in the case of the second and third columns, the number of RMSE smaller than those of a random walk. Results are shown for forecast horizons of $h = 1, 6, 12$. The currency abbreviations stand for Australian Dollar (AUD), Brazilian Real (BRL), Canadian Dollar (CAD), Danish Krone (DKK), Indian Rupee (INR), Mexican Peso (MXN), New Zealand Dollar (NZD), Norwegian Krone (NOK), South African Rand (ZAR), Swedish Krona (SEK), Swiss Franc (CHF), British Pound (GBP), and Euro (EUR).

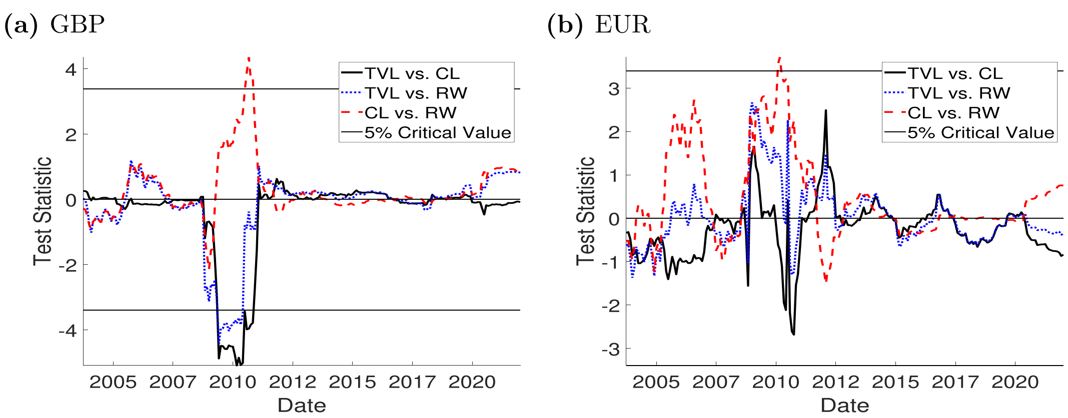

FIGURE 9 Fluctuation test, $h = 1$. *Note:* This figure plots the results of the Giacomini and Rossi (2010) fluctuation test. The solid blue line is the test statistic, the dotted red lines are the 5% critical values. The results are based on a rolling window of $m = 20$ and a quadratic loss function.

rejects for the JPY, however using an absolute loss function, it rejects for the AUD (i.e., in favour of the TVL model). On the other hand, the tests reject in favour of the random walk throughout when comparing it to the CL model. All in all, the results demonstrate that time-varying loadings improve the fit of the model, especially for one-step-ahead forecasts. In a companion working paper (Hillebrand et al., 2020), we conducted the out-of-sample estimation using the in-sample dataset, that is, without real-time data. The TVL model also consistently outperformed the CL model, albeit with fewer rejections overall. We conjecture that the Kalman filter is able to filter out noise in the real-time data, which led to an even better performance of the TVL relative to the CL model.

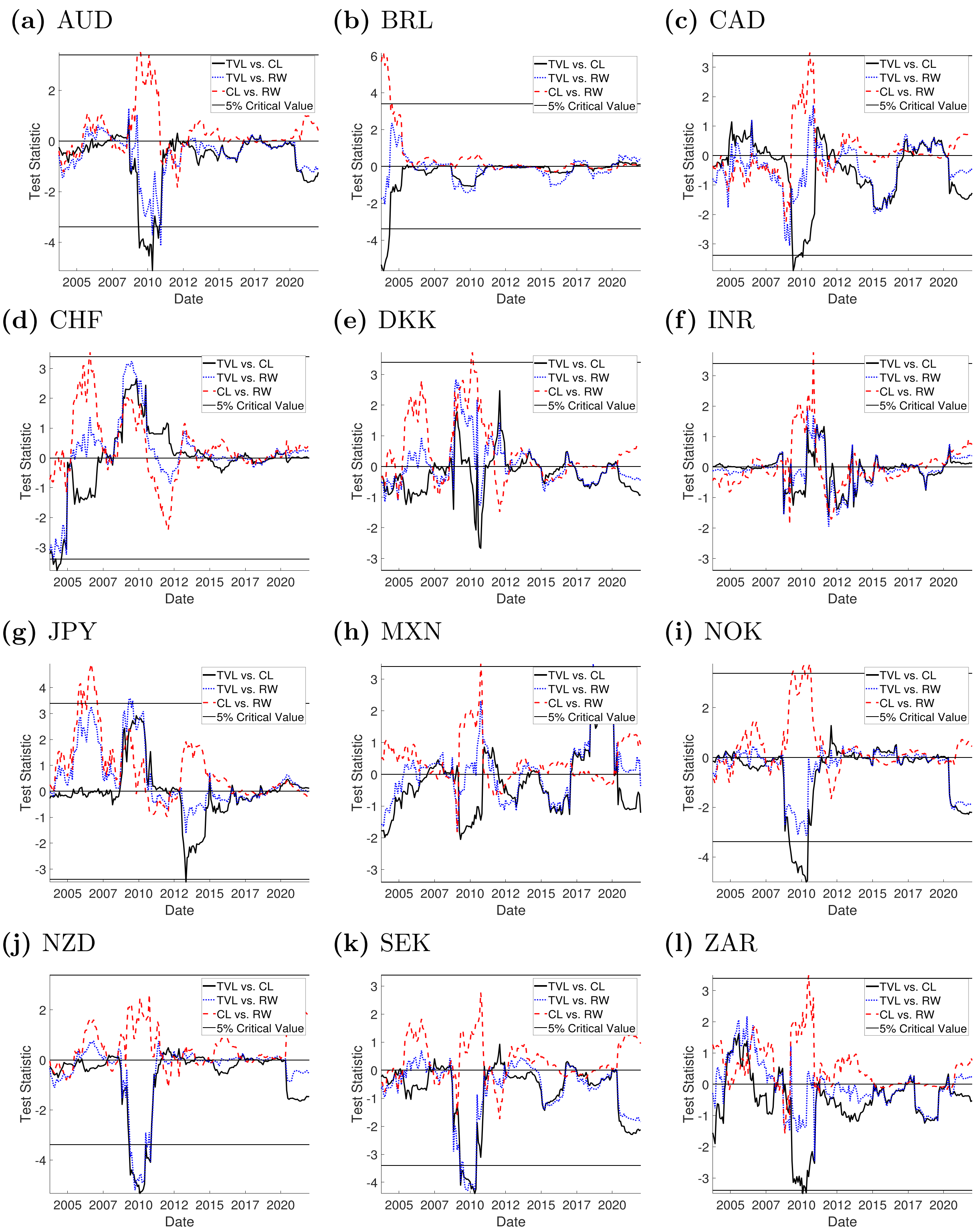

FIGURE 10 Fluctuation test, $h = 1$. *Note:* This figure plots the results of the Giacomini and Rossi (2010) fluctuation test. The solid blue line is the test statistic, the dotted red lines are the 5% critical values. The results are based on a rolling window of $m = 20$ and a quadratic loss function.

In addition to selecting a model based on relative global predictive ability, it is interesting to examine how the relative predictive ability of two models changes over time. Figure 9 plots the results of the Giacomini and Rossi (2010) fluctuation test for GBP and EUR at each point in time for $h = 1$.[13] The solid black line represents the test statistic for the TVL compared to the CL model, the dotted blue line the statistic for the TVL model against a random walk, and the dashed red line corresponds to the test statistic for the CL model against the random walk. The horizontal lines are the 5% critical values. The sign of the test statistic corresponds to the sign of the MSE. When the test statistic falls below the negative critical value, the test rejects the null hypothesis of equal predictive ability in favour of the TVL model (or the CL model against the random walk). The results are obtained using a rolling window of $m = 20$ and a quadratic loss function. Although the Giacomini and White (2006) test does not provide evidence that the time-varying loadings model is globally more accurate in case of the GBP (see Table 3), the fluctuation test rejects in favour of the TVL model both against the CL model and the random walk during the financial crisis. Figure 10 shows the results for the remaining currencies. The fluctuation test rejects the null hypothesis that TVL and CL have equal predictive ability for 10 currencies (AUD, BRL, CAD, CHF, GBP, JPY, NOK, NZD, SEK, and ZAR). The Giacomini and White (2006) test only rejects the null hypothesis of equal predictive ability for 4 currencies (BRL, NOK, NZD, and SEK, Table 3). That is, for an additional six currencies, we find evidence that accounting for instabilities in factor loadings leads to improved local predictive ability. In most cases, these pockets of predictability occur during the financial crisis, which indicates that the TVL model performed particularly well in that period. The fluctuation test statistic never attains the positive critical value, meaning the null hypothesis is never rejected in favour of the constant loadings model. The test further indicates that the random walk has superior predictive ability when compared against the CL model in all except two cases (NZD and SEK). In contrast, the fluctuation test only rejects twice in favour of the random walk against the TVL model (JPY and MXN) and five times in favour of the TVL model (AUD, CHF, GBP, NZD, and SEK). On closer examination, the CL model performs particularly poorly against the random walk during periods in which the TVL model does well, especially for AUD, BRL, CAD, and GBP. For these 4 currencies, the random walk significantly outperforms the CL model during periods in which the TVL model beats the random walk.

7 | CONCLUSIONS

In this paper, we studied the unstable relationship between exchange rates and macroeconomic factors. Using a novel econometric approach, proposed in Mikkelsen et al. (2019), we showed that allowing for time-varying factor loadings increases the percentage of explained variation in exchanges rates by an order of magnitude. In addition, taking the aforementioned instabilities into consideration improves the relative out-of-sample predictive ability of the model globally, and yields better forecast of sign changes in exchange rates.

We extracted macroeconomic fundamentals as principal components from a new dataset that combines all 272 vintage releases of the FRED-MD database and a large number of variables sourced from the OECD. The unobserved time-varying loadings are estimated using a recently proposed two-step maximum likelihood estimator for high-dimensional factor models (Mikkelsen et al., 2019). The model is applied to 14 currencies vis-à-vis the US Dollar and the results show that failure to account for the instabilities between exchange rates and fundamentals is by no means innocuous. In-sample, the time-varying loadings model achieves a median R^2 of 52% percentage points compared to 2% for the constant loadings benchmark. We showed that out-of-sample, the time-varying loadings model exhibits significantly better forecast accuracy. This holds true both when comparing their predictive ability directly, and in terms of the improvements the time-varying model generates relative to a random walk. When evaluating the forecasts individually, the time-varying loadings model outperforms the constant loadings model at predicting directional exchange rate changes. To consider potentially unstable forecasting performance, we evaluated the relative predictive accuracy of the forecasts using the Giacomini and Rossi (2010) fluctuation test. In addition to higher global forecast accuracy, time-varying loadings improved forecasts locally around the financial crisis. This paper provides strong evidence that the relationship between macroeconomic fundamentals and exchange rates is highly unstable.

[13] Plots for $h = 6{,}12$ are reported in Appendix B.

ACKNOWLEDGEMENTS

We thank the editor, Barbara Rossi, and two anonymous referees for their invaluable comments and suggestions, which led to an improved version of the paper. We are grateful to Roy Batchelor, Ian Marsh, Kate Phylaktis and Lucio Sarno for useful comments. We thank seminar participants at Aarhus University, the University of Maastricht, the Tinbergen Institute, and Oxford University for very helpful comments and suggestions on a previous version of the paper. The usual disclaimer applies. We also acknowledge financial support from the Centre for Econometric Analysis at Bayes Business School.

OPEN RESEARCH BADGES

This article has been awarded Open Data Badge for making publicly available the digitally-shareable data necessary to reproduce the reported results.

DATA AVAILABILITY STATEMENT

Data and code for replication have been uploaded to the ZBW Journal Data Archive.

REFERENCES

Ahn, S. C., & Horenstein, A. R. (2013). Eigenvalue ratio test for the number of factors. *Econometrica*, *81*(3), 1203–1227.

Alessi, L., Barigozzi, M., & Capasso, M. (2010). Improved penalization for determining the number of factors in approximate factor models. *Statistics & Probability Letters*, *80*(23-24), 1806–1813.

Bacchetta, P., & van Wincoop, E. (2004). A scapegoat model of exchange-rate fluctuations. *American Economic Review P&P*, *94*(2), 114–118.

Bacchetta, P., & van Wincoop, E. (2009). On the unstable relationship between exchange rates and macroeconomic fundamentals. (*w15008*): NBER Working Paper.

Bacchetta, P., & van Wincoop, E. (2012). Modeling exchange rates with incomplete information. In James, J., Marsh, I. W., & Sarno, L. (Eds.), *Handbook of exchange rates* (pp. 375–390) John Wiley & Sons, Inc.

Bacchetta, P., & van Wincoop, E. (2013). On the unstable relationship between exchange rates and macroeconomic fundamentals. *Journal of International Economics*, *91*(1), 18–26.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, *71*(1), 135–171.

Bai, J., & Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, *70*(1), 191–221.

Barigozzi, M., Hallin, M., Soccorsi, S., & von Sachs, R. (2021). Time-varying general dynamic factor models and the measurement of financial connectedness. *Journal of Econometrics*, *222*(1), 324–343.

Bates, B. J., Plagborg-Møller, M., Stock, J. H., & Watson, M. W. (2013). Consistent factor estimation in dynamic factor models with structural instability. *Journal of Econometrics*, *177*(2), 289–304.

Bekiros, S. D. (2014). Exchange rates and fundamentals: Co-movement, long-run relationships and short-run dynamics. *Journal of Banking & Finance*, *39*, 117–134.

Boivin, J., & Ng, S. (2005). Understanding and comparing factor-based forecasts. *International Journal of Central Banking*, *1*(3), 117–151.

Byrne, J. P., Korobilis, D., & Ribeiro, P. J. (2018). On the sources of uncertainty in exchange rate predictability. *International Economic Review*, *59*(1), 329–357.

Caruso, A., & Coroneo, L. (2022). Does real-time macroeconomic information help to predict interest rates? *Journal of Money, Credit, and Banking*. forthcoming.

Choi, I., & Jeong, H. (2019). Model selection for factor analysis: Some new criteria and performance comparisons. *Econometric Reviews*, *38*(6), 577–596.

Clark, T. E., & McCracken, M. W. (2001). Test of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics*, *105*, 85–110.

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, *13*(3), 253–265.

Doz, C., Giannone, D., & Reichlin, L. (2011). A two-step estimator for large approximate dynamic factor models based on Kalman filtering. *Journal of Econometrics*, *164*, 188–205.

Doz, C., Giannone, D., & Reichlin, L. (2012). A quasi-maximum likelihood approach for large, approximate dynamic factor models. *The Review of Economics and Statistics*, *94*(4), 1014–1024.

Engel, C. (2014). Exchange rates and interest parity. In Gopinath, G., Helpman, E., & Rogoff, K. (Eds.), *Handbook of international economics* (Vol. *4*, pp. 453–522). Elsevier.

Engel, C., Mark, N. C., & West, K. D. (2015). Factor model forecasts of exchange rates. *Econometric Reviews*, *34*(1-2), 32–55.

Engel, C., & West, K. D. (2005). Exchange rates and fundamentals. *Journal of Political Economy*, *113*(3), 485–517.

Fratzscher, M., Rime, D., Sarno, L., & Zinna, G. (2015). The scapegoat theory of exchange rates: The first tests. *Journal of Monetary Economics*, *70*, 1–21.

Giacomini, R., & Rossi, B. (2010). Forecast comparisons in unstable environments. *Journal of Applied Econometrics*, *25*(4), 595–620.

Giacomini, R., & White, H. (2006). Tests of conditional predictive ability. *Econometrica*, *74*(6), 1545–1578.

Hallin, M., & Liška, R. (2007). Determining the number of factors in the general dynamic factor model. *Journal of the American Statistical Association*, *102*(478), 603–617.

Hillebrand, E., Mikkelsen, J. G., Spreng, L., & Urga, G. (2020). *Exchange rates and macroeconomic fundamentals: Evidence of instabilities from time-varying factor loadings*: Department of Economics and Business Economics, Aarhus University.

Kapetanios, G. (2010). A testing procedure for determining the number of factors in approximate factor models with large datasets. *Journal of Business & Economic Statistics*, *28*(3), 397–409.

Kouwenberg, R., Markiewicz, A., Verhoeks, R., & Zwinkels, R. C. J. (2017). Model uncertainty and exchange rate forecasting. *Journal of Financial and Quantitative Analysis*, *52*(1), 341–363.

Leitch, G., & Tanner, J. E. (1991). Economic forecast evaluation: Profits versus the conventional error measures. *American Economic Review*, *81*(3), 580–590.

McCracken, M. W., & Ng, S. (2016). FRED-MD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, *34*(4), 574–589.

Meese, R., & Rogoff, K. (1983). Empirical exchange rate models of the seventies. *Journal of International Economics*, *14*, 3–24.

Mikkelsen, J. G., Hillebrand, E., & Urga, G. (2019). Consistent estimation of time-varying loadings in high-dimensional factor models. *Journal of Econometrics*, *208*(2), 535–562.

Obstfeld, M., & Rogoff, K. (2000). The six major puzzles in international macroeconomics: Is there a common cause? *NBER Macroeconomics Annual*, *15*, 339–412. http://www.nber.org/books/bern01-1

Pesaran, M. H., & Timmermann, A. (1992). A simple nonparametric test of predictive performance. *Journal of Business & Economic Statistics*, *10*(4), 461–465. http://www.jstor.com/stable/1391822

Pozzi, L., & Sadaba, B. (2020). Detecting scapegoat effects in the relationship between exchange rates and macroeconomic fundamentals: A new approach. *Macroeconomic Dynamics*, *24*(4), 951–994.

Rossi, B. (2006). Are exchange rates realls random walks? Some evidence robust to parameter instability. *Macroeconomic Dynamics*, *10*, 2038.

Rossi, B. (2013). Exchange rate predictability. *Journal of Economic Literature*, *51*(4), 1063–1119.

Rossi, B. (2021). Forecasting in the presence of instabilities: How do we know whether economic models work and how to improve them. *Journal of Economic Literature*, *59*(4), 1135–1190.

Sarno, L., & Valente, G. (2009). Exchange rates and fundamentals: Footloose or evloving relationship. *Journal of the European Economic Association*, *7*(4), 786–830.

Shumway, R. H., & Stoffer, D. S. (1982). An approach to time series modeling and forecasting using the EM algorithm. *Journal of Time Series Analysis*, *3*, 253–264.

Su, L., & Wang, X. (2017). On time-varying factor models: Estimation and testing. *Journal of Econometrics*, *198*(1), 84–101.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of the article.

How to cite this article: Hillebrand E., Mikkelsen J. G., Spreng L., & Urga G. (2023). Exchange rates and macroeconomic fundamentals: Evidence of instabilities from time-varying factor loadings. *Journal Applied of Econometrics*, 1–21. https://doi.org/10.1002/jae.2984