Imperfect Exchange Rate Expectations*

Giacomo Candian† HEC Montréal

Pierre De Leo‡ University of Maryland

First draft: December 30, 2020
This draft: October 7, 2021

Abstract

Using survey data, we document that predictable exchange rate forecast errors are responsible for the uncovered-interest-parity (UIP) puzzle and its reversal at longer horizons. We develop a general-equilibrium model based on shock misperception and over-extrapolative beliefs that reconciles these and other major exchange rate puzzles. These beliefs distortions generate both under- and overreaction of expectations that account for the predictability of forecast errors about interest rates, exchange rates, and other macroeconomic indicators. In the model, forecast errors are endogenous to monetary policy and explain the change in the behavior of UIP deviations that emerged after the global financial crisis.

Keywords: Exchange Rate Forecast Errors; Beliefs Formation; UIP Violations; New Fama Puzzle.
JEL Classification: E43; E7; F31.

*We thank Luca Gemmi, Nils Gornemann, Sebnem Kalemli-Ozcan, Alexandre Kohlhas, Ricardo Reis, Vania Stavrakeva, Luminita Stevens and Liliana Varela for useful discussions, as well as seminar and conference participants at HEC Montreal, CEPR & IMF 6th Conference, Irish Economic Association Annual Conference, AFES 2021, NASMES 2021, AMES 2021 and ESAM 2021. Thanks also to Wenbo Yu for his expert research assistance. Candian gratefully acknowledges financial support from the Insight Development Grant of the Social Sciences and Humanities Research Council.

†Department of Applied Economics, 3000, Chemin de la Côte-Sainte-Catherine, Montréal QC, H3T 2A7 Canada. Telephone: 514-340-6810. Email: giacomo.candian@hec.ca.

‡Department of Economics, 4118D Tydings Hall, College Park, MD, 20742, United States. Email: deleop@umd.edu.

Uncovered interest rate parity (UIP) predicts that an increase in the domestic interest rate should coexist with an expected depreciation of the domestic currency, so as to eliminate any excess return from holding a domestic bond instead of a foreign one. In the data, however, high interest rate currencies tend to appreciate in the short run and carry, therefore, a positive excess return. This empirical regularity is known as the UIP puzzle. Furthermore, recent evidence revealed that the puzzle is more complex: high interest rates today predict a negative excess return 3 to 6 years into the future. That is, the UIP puzzle reverses directions at longer horizons.[1]

We begin by showing that the classic UIP puzzle and its subsequent reversal are *both* accounted for by systematic forecast errors. Using survey data on exchange rate forecasts for advanced economies, we confirm earlier evidence that high interest rate differentials are systematically associated with a belief that the exchange rate will depreciate in the short-run but are followed, on average, by an appreciation. This under-prediction of the value of the domestic currency explains its positive short-run excess returns.[2] In addition, we document a new fact: high interest rates today are systematically associated with a future over-prediction of the value of the domestic currency. That is, the domestic currency eventually depreciates faster than expected, accounting for the negative excess returns at longer horizons. Quantitatively, these predictable exchange rate forecast errors closely track the UIP deviations at short and longer horizons, suggesting that systematic expectation errors, as opposed time-varying expected excess returns, are responsible for both puzzles.

In this paper we propose a mechanism based on distorted beliefs that explains the UIP puzzle and its reversal. Beliefs in our model exhibit an initial under-reaction and a subsequent over-reaction to changes in fundamentals. When placed in a general equilibrium environment, our mechanism reproduces the well-documented predictability of forecast errors about exchange rates *and* key macroeconomic indicators, and accounts for several exchange-rate puzzles.[3] These forecast errors are endogenous to policy and explain the change in the behavior of UIP deviations that emerged after the

[1] Classic references to the UIP puzzle are Bilson (1981) and Fama (1984). An excellent survey on the more recent literature is Engel (2014). Evidence on the UIP puzzle reversal at longer horizons can be found in Bacchetta and Van Wincoop (2010), Engel (2016), and Valchev (2020).

[2] Froot and Frankel (1989) show that forecast errors account for the classic UIP puzzle for advanced economies. Kalemli-Ozcan and Varela (2021) confirm that forecast errors explain the UIP puzzle for advanced economies, but they show that this is not the case for emerging economies.

[3] That is, our model speaks to recent evidence on the properties of survey expectations for both macro variables (Coibion and Gorodnichenko, 2015; Angeletos et al., 2020; Kohlhas and Walther, 2021) and exchange rates (Stavrakeva and Tang, 2020b; Bussiere et al., 2018).

global financial crisis. Our findings point to an important role of beliefs distortions in explaining exchange rate dynamics.

To illustrate the essence of our mechanism, we first formulate a simple model of distorted beliefs about the future interest rate path. We specify an environment in which investors have to determine whether interest rate shocks are transitory or persistent. As in Gourinchas and Tornell (2004), investors overstate the relative importance of transitory interest rate shocks. In addition, we assume that investors over-extrapolate, that is, they perceive persistent shocks to be more autocorrelated than they actually are. The exchange rate is then determined by UIP under investors' subjective beliefs about the interest rate path.

Our mechanism features both short-run under-reaction and longer-run over-reaction to changes in the interest rate differential. To gain intuition, suppose that the U.S. interest rate rises and then gradually returns to its initial level. If expectations were rational, the US dollar would appreciate immediately and gradually return to its long-run value. Instead, expectations in our model feature two biases which dominate at different horizons. First, investors misperceive the shock to be partly transitory and initially under-react to the interest rate increase. When the interest rate turns out to be higher than first expected, investors' upward revision of the future interest rate path causes an appreciation of the dollar, which experiences positive excess returns in the short run. Second, over-extrapolation leads investors to perceive shocks to be more persistent than they actually are, causing the dollar to be, at some point, over-appreciated relative to rational expectations. This excessive dollar appreciation entails a subsequent downward revision of the interest rate path that generates negative excess returns in the medium run. This mechanism thus explains the UIP puzzle and its reversal with forecast errors rather than relying on a source of expected excess returns such as risk premia.

We then provide empirical evidence in support of the mechanism. In our theory of distorted beliefs, exchange rate forecast errors arise from errors in interest rates forecasts: interest rate expectations under-react in the short run while they over-react in the longer run to an interest rate innovation.[4] This is a necessary condition to generate the reversal in exchange rate forecast errors. We verify this key prediction

[4]We do not argue that exchange rate forecast errors arise solely from interest rates forecast errors. Indeed, the mechanism that we propose would resolve both puzzles even if similarly distorted beliefs were held on other exchange rate determinants, as long as these determinants comove with interest rate differentials in the right way. Nevertheless, distorted beliefs about interest rates are a natural starting point from asset-view models of exchange rate determination.

using data on forecasts of interest rates (not examined in our analysis thus far) and find that it is empirically supported: interest rate expectations are excessively pessimistic in the short run while excessively optimistic in the longer-run, favoring our proposed interpretation of the UIP puzzle and subsequent reversal.

We then take all these insights to a two-country general-equilibrium (GE) environment where we can study the broader macroeconomic impact of these belief distortions. The key departure from a standard model is that agents misperceive and over-extrapolate shocks to fundamentals. Given these beliefs, macroeconomic outcomes, including interest rates, and expectations thereof are determined endogenously. We discipline the extent of belief distortions with recent evidence on forecast errors for U.S. output and inflation (Coibion and Gorodnichenko, 2015; Angeletos et al., 2020; Kohlhas and Walther, 2021). While belief distortions are disciplined by forecast errors about domestic macro variables, the model successfully reproduces several moments of exchange rate forecast errors, including their predictability with past exchange rate changes (Stavrakeva and Tang, 2020b). We find this remarkable because it suggests that expectations about very different macro quantities and prices, typically studied in isolation, may be subject to similar behavioral distortions.

The GE model delivers several insights and testable predictions. First, it rationalizes various open-economy puzzles, including the dual UIP puzzle, the PPP puzzle, and the Backus-Smith puzzle. Second, the UIP puzzle and its reversal in GE arise under both supply and demand shocks typically considered in open-economy models. Third, when allowing for both belief distortions and time-varying liquidity premia, the GE model reproduces the empirical predictability of both *ex-ante* and *ex-post* excess currency returns.

Finally, a distinctive feature of our GE model is that the dynamics of forecast errors (and, therefore, of UIP violations) critically depend on agents' biases *and* the structural relationships implied by the model. In particular, the dynamic properties of interest rates and ensuing forecast errors are endogenous to policy. This insight allows to confront our model with recent evidence about short-run UIP deviations. Specifically, Bussiere et al. (2018) show that high interest rate currencies have been earning a negative short-run excess return following the global financial crisis – in contrast to the positive excess returns they earned pre-2008. Our model naturally reproduces this switch in the direction of excess return predictability under a monetary policy regime that does not satisfy the Taylor principle (to characterize the later sample, as in Bianchi and Melosi, 2017).[5] As in the data, this switch in excess returns is explained by a change in the comovement between forecast errors and interest differentials.

Relation to the literature A long tradition in macroeconomics and finance uses survey data to inform theories of expectations formation (eg., Coibion and Gorodnichenko, 2015; Bordalo et al., 2020). In the foreign exchange market, Frankel and Froot (1987) use survey expectations to test alternative theories of expectations formation, while Stavrakeva and Tang (2020b) document that forecast errors can be predicted by past appreciations, and link foreign exchange derivatives positions to expected exchange rates. Engel et al. (2008) and Stavrakeva and Tang (2020a) study whether exchange rates respond to survey-based measures of macroeconomic news. Froot and Frankel (1989) use survey data to disentangle forecast errors from expected excess returns in explaining the classic UIP puzzle, and attribute the short-run UIP deviations mainly to forecast errors.[6] Kalemli-Ozcan and Varela (2021) document several properties of exchange rates and forecast errors in both advanced and emerging economies, and highlight the role of policy uncertainty in driving survey-based UIP premia. Our empirical contribution is to document that in advanced economies predictable exchange rate forecast errors closely track UIP deviations at *all* horizons. We argue that a theory of exchange rates that explains excess returns must be consistent with these facts about exchange rate beliefs.

Our findings reveal that forecast errors account for both the classic UIP puzzle (Fama, 1984) and for its reversal at longer horizons documented by Bacchetta and Van Wincoop (2010), Engel (2016) and Valchev (2020). Available explanations of the non-monotonic patterns of excess returns predictability rely on rational expectations: Engel (2016) and Valchev (2020) advance theories based on time-varying convenience yields, while Bacchetta and van Wincoop (2021) propose an explanation based on gradual portfolio adjustment.[7,8] By contrast, we put forward a theory of UIP deviations

[5]In this regime the unique equilibrium is determined by an "active" fiscal rule, as in Leeper (1991).

[6]Chinn and Frankel (2019) confirm this finding in recent samples. Bacchetta et al. (2009) show that the short-run predictability of excess returns and forecast errors extends to other financial markets, such as stock markets, bond markets and money markets in various countries.

[7]See also Dahlquist and Pénasse (2021), Chernov and Creal (2020), and Jiang et al. (2021).

[8]Other rational expectations explanations of the classic UIP puzzle are based on imperfect financial markets (Gabaix and Maggiori, 2015) or time-varying risk due to rare disasters (Farhi and Gabaix, 2015), long-run risk (Bansal and Shaliastovich, 2012; Colacito and Croce, 2013), habits (Verdelhan, 2010), or a time-varying degree of asset market segmentation (Alvarez et al., 2009).

based on distorted beliefs that reconciles our evidence on exchange rate expectations.[9]

Our evidence on forecast error predictability implies that departures from full information alone cannot account for observed UIP violations if expectations are rational and interest rates are observed. In fact, models that rely exclusively on imperfectly observed fundamentals cannot generate predictable UIP deviations (e.g., Candian, 2019; Hetting et al., 2021). Abandoning the rational expectation paradigm inevitably results in behavioral biases that lead to either under-reaction or over-reaction. Models with misperceptions about transitory shocks such as Gourinchas and Tornell (2004) generate under-reaction of exchange rates to changes in interest rate differentials that explain the UIP puzzle but miss a force of magnified adjustment that accounts for the subsequent reversal. Any model of overconfidence of private signals (e.g., Burnside et al., 2011) also faces the same limitation.[10] Conversely, over-extrapolative beliefs generate overreaction to current disturbances but this bias alone leads to short-run UIP deviations of the wrong sign, as we show in our analysis.[11] Our contribution with respect to this literature is to identify which of these behavioral biases are needed to explain the the predictability reversal and, most importantly, to build a framework where these biases have GE effects that can explain a wider range of exchange rate phenomena.

The reversal of exchange rate forecast errors at longer horizons echoes recent evidence of Angeletos et al. (2020), who find a similar pattern in U.S. unemployment and inflation expectations following business cycle shocks. The cyclical pattern in the predictability of forecast errors is thus a pervasive phenomenon. A contribution of this paper is to show that a simple set of belief distortions, when placed in a GE environment, simultaneously accounts for many well-know facts of forecast errors for U.S. output, inflation, interest rates as well as interest rate differentials and exchange rates.

Our GE model belongs to the New Open Economy Macroeconomics literature initiated by Obstfeld and Rogoff (1995), and reproduces several classic and new exchange rate-related puzzles. Recently, Itskhoki and Mukhin (2021) show that introducing a time-varying wedge in the UIP condition resolves a variety of exchange rate puzzles in

[9]Bunsupha (2018) postulates a form of non-rational exchange rate expectations that react to past appreciations. These expectations result in an oscillatory behavior of exchange rate forecast errors that is inconsistent with our empirical evidence.

[10]Explanations based on exogenous noise trader shocks are also unable to generate a reversal in excess returns forecast errors (see, e.g., Bacchetta and Van Wincoop, 2006; Jeanne and Rose, 2002; Itskhoki and Mukhin, 2021).

[11]Over-extrapolative beliefs have been used in the finance literature to explain a variety of asset-price phenomena (Barberis et al., 1998; Daniel et al., 1998; Hong and Stein, 1999; Greenwood and Shleifer, 2014; Gennaioli et al., 2016; Guo and Wachter, 2019).

international macro. Distorted beliefs generate a UIP wedge in our model that is due to expectations' under and over-reaction to changes in fundamentals. This wedge is endogenous and indeed helps explain the moments of exchange rates examined in that paper, but also other ones including the excess comovement puzzle (Engel, 2016) and the sign change in the Fama regression coefficient occurred after the global financial crisis (Bussiere et al., 2018).

In parallel and independent work, Valente et al. (2021) develop a model with noisy private signals and extrapolation about an exogenous interest rate process, and show that such a model explains several exchange rate puzzles. Our paper and theirs bring distinct, but complementary, empirical evidence in support of the mechanism underlying both models. Using a different form of departure from rationality, constraints on the complexity of investors' statistical models, Molavi et al. (2021) generate patterns that are simultaneously consistent with the UIP puzzle and its subsequent reversal. An important distinction that sets us apart from these two papers is that we embed distorted beliefs into a full-fledged two-country GE model in which monetary policy is endogenous. This quantitative framework allows us to confront a wide set of exchange rate and macro moments from the model with the data, compare the properties of forecast errors of several macro variables within a unique framework, and study the dynamics of UIP deviations in different monetary regimes. Beside being quantitatively successful in reproducing several open economy moments, a unique takeaway of our model is that a common set of beliefs distortions jointly accounts for the predictability of macro and exchange rate forecast errors and rationalizes the post-2008 sign change in UIP deviations with a change in monetary policy.

1 Excess Returns and Forecast Errors

In this section, we first reproduce the empirical evidence on the predictability of excess currency returns at short and long horizons following the approach in Valchev (2020). We then isolate the portion of this predictability due to forecast errors.

We use monthly data on interest rates and spot bilateral exchange rates against the US dollar for 8 advanced economies. Appendix A provides a detailed description of the data. Our baseline sample period is 1990:M1-2007:M12, while Appendix F reports our findings for a longer sample (1990:M1-2019:M7).[12]

[12] Our baseline sample ends in 2007:M12 because of plausible structural breaks in the driving pro-

Our analysis considers a home (the US) and a foreign country. We denote with r_t the log of the one-period nominal interest rate on period-t deposits that pay off in period $t+1$ and with r_t^* the corresponding foreign interest rate. Let s_t be the log of the foreign exchange rate, the price of a unit of foreign currency expressed in US dollars. A rise in s_t thus denotes a depreciation of the domestic currency. We use Λ_{t+1} to denote the excess return on a foreign interest rate deposit between periods t and $t+1$, inclusive of foreign currency appreciation over the same period; that is:

$$\Lambda_{t+1} \equiv s_{t+1} - s_t - x_t,$$

where $x_t^j = r_t - r_t^*$ is the nominal interest rate differential.

1.1 Excess currency return predictability

Under UIP the expected exchange rate depreciation, $\mathrm{E}_t\, \Delta s_{t+1}$, offsets any potential gap in interest rates, x_t, such that there are no arbitrage opportunities:

$$\mathrm{E}_t\, \Delta s_{t+1} = x_t, \tag{1}$$

where we denote with E_t the expectation conditional on time-t information set.

Thus, under UIP, next period excess currency returns, Λ_{t+1}, should be unforecastable. If expectations are rational, we can test this hypothesis by estimating the regression equation:

$$\Lambda_{t+1} = \alpha + \beta_1 x_t + \varepsilon_{t+1}. \tag{2}$$

Under the joint null hypothesis of rational expectations and UIP, $\beta_1 = 0$, so that excess returns are, on average, not forecastable, i.e., $\mathrm{E}_t\, \Lambda_{t+1} = 0$. In contrast to this prediction, many papers have documented that excess returns are predictable by estimating $\beta_1 < 0$, which means that countries with a positive interest rate differential, on average, earn positive excess returns on their domestic currency (equivalently negative excess returns on the foreign currency). A $\beta_1 < 0$ represents the UIP puzzle, also known as Fama puzzle.

Recent work by Engel (2016) and Valchev (2020) has documented that interest rates predict excess currency returns not only in the near future but also at longer horizons. From the standpoint of the UIP condition, this is equally puzzling. Indeed

cesses for interest rates and exchange rates during the period of the global financial crisis.

by rolling equation (1) forward k periods and applying the law of iterated expectations one obtains that, for any $k \geq 1$:

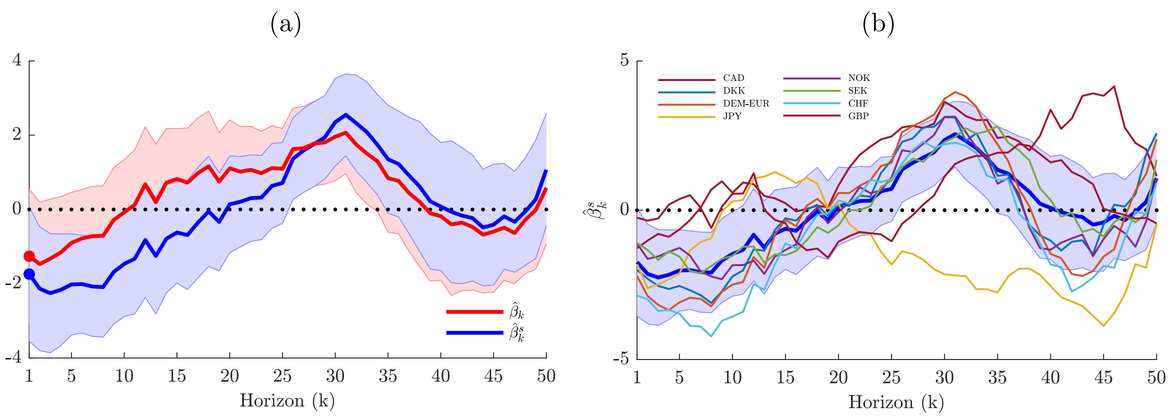
$$\mathrm{E}_t \, \Lambda_{t+k} = 0,$$

which means that according to UIP any future one-period excess currency returns should be unforecastable. To test these conditions, we follow Engel (2016) and Valchev (2020) and estimate the following panel regression equation:

$$\Lambda_{j,t+k\times 3} = \alpha_{j,k} + \beta_k x_t^j + \varepsilon_{j,t+k}, \tag{3}$$

where the index j represents a currency and $x_t^j = r_t - r_t^j$.[13] UIP and rational expectations imply $\beta_k = 0$ for any $k \geq 1$, with the original Fama (1984) regression being the case of $k = 1$.

Figure 1: UIP Regression Results (1990:M1-2007:M12)



Notes: Panel (1a) reports the panel estimates of β_k and β_k^s in equations (3) and (7), respectively, along with their 95% confidence interval. Panel (1b) reports the country-by-country estimates of β_k^s, along with the panel counterpart. Each horizon is a quarter.

Figure 1a depicts the estimated $\hat{\beta}_k$ in our dataset. The red line represents the point estimates of β_k at different horizons $k \geq 1$, while the red shaded area represents 95% confidence interval around each estimate of β_j. These are panel estimates and the standard errors are computed following Driscoll and Kraay (1998) to correct for heteroskedasticity, and both serial and cross-equation correlation. The estimates

[13] The variable $\Lambda_{j,t+k\times 3}$ is the one-quarter excess return on the jth currency from month $t+(k-1)\times 3$ to $t + k \times 3$. We use this investment maturity to be consistent with the 3-month forecast errors data that we use in Section 1.2.

confirms the classic result that high interest rates forecast negative next-period excess foreign-currency returns, as captured by the negative estimate of β_1. Moreover, high interest rates predict negative one-period excess return up to 20 quarters into the future, before changing sign and predicting positive excess returns up to about 50 quarters ahead. These patterns of excess return predictability align well with those found by Engel (2016) and Valchev (2020). Virtually identical results obtain in the longer sample 1990:M1-2019:M7 (see Figure A.1a).

1.2 Decomposing the Excess Return Coefficient

The above analysis documents the presence of predictable excess returns at short and long horizons but it is silent on their origins. Are the excess returns the consequence of a risk premium or are they the reflection of systematic forecast errors? Regression (3) does not allow us to answer this question directly because it implicitly imposes that expectations are rational, instead of measuring expectations directly.

For this reason, we follow the approach of Froot and Frankel (1989) and use a direct measure of exchange rate expectations to disentangle forecast errors from expected excess currency returns, ξ_t, defined as:

$$\xi_t = x_t - (\mathrm{E}_t^s\, s_{t+1} - s_t). \tag{4}$$

Here, E_t^s denotes subjective expectations which need not be rational. Equation (4) states that if the domestic interest rate differential exceeds the subjective expected depreciation of the domestic currency, then the domestic currency carries a positive expected excess currency return. For brevity, we refer to ξ_t as risk premia while acknowledging that these ex-ante excess returns may be driven by other factors.

Using equation (4) we can rewrite the definition of one-period excess currency returns at time $t+k$ as:

$$\Lambda_{t+k} = (s_{t+k} - \mathrm{E}_{t+k-1}^s\, s_{t+k}) - \xi_{t+k-1}, \tag{5}$$

which states that ex-post excess returns can arise either because of forecast errors or due to a risk premium. By taking the covariance of both sides of equation (5) with the interest rate differential x_t and dividing by its variance we obtain:

$$\underbrace{\frac{\text{Cov}\left(\Lambda_{t+k}, x_t\right)}{\text{Var}(x_t)}}_{\beta_k} = \underbrace{\frac{\text{Cov}\left(s_{t+k} - \text{E}^s_{t+k-1}\, s_{t+k}, x_t\right)}{\text{Var}(x_t)}}_{\beta^s_k} - \underbrace{\frac{\text{Cov}\left(\xi_{t+k-1}, x_t\right)}{\text{Var}(x_t)}}_{\beta^\xi_k}. \tag{6}$$

This decomposition suggests that to understand whether the predicability of excess currency returns is due to forecast errors or risk premia we can regress a measure of forecast errors from the survey data on the interest rate differential:

$$s_{t+k\times 3} - \text{E}^s_{t+(k-1)\times 3}\, s_{t+k\times 3} = \gamma + \beta^s_k x_t + \eta_{t+k}, \tag{7}$$

and then compare the estimates of β_k and β^s_k. If β^s_k is close to β_k then predictability of excess currency returns is due to the predicability of forecast errors.

Our data on exchange rate forecasts for regression (7) come from *Consensus Economics* and are described in Appendix A. The point estimates and confidence intervals of β^s_k are depicted in Figure 1 in blue. We emphasize two key findings. First, interest rates differentials can predict forecast errors about exchange rates. In particular, a high interest rate differential is associated with a systematic negative forecast error of the exchange rate in the short run and a positive one in the medium run. When the interest rate differential is high, agents underpredict the rate of exchange rate appreciation in the short run, while they overpredict it in the longer run.

The second main takeaway is that the predictability of forecast errors largely accounts for the predictability of excess returns. This is apparent from the fact that the red line closely tracks the blue line, and generally lies in the 95% confidence interval of the latter. In other words, high interest rate differentials predict negative short-run excess returns on the foreign currency because high interest rate differential are associated with a belief that the exchange rate will appreciate less than it actually does, and conversely in the longer run. These patterns are consistent with Kalemli-Ozcan and Varela's (2021) evidence that survey-based UIP premia are largely unpredictable by interest rates at all horizons in advanced economies. Similar results obtain in individual countries' regressions (see Figure 1b) as well as in the longer sample 1990:M1-2019:M7 (see Figure A.1).

2 A Simple Model of Distorted Beliefs

We now describe a model of exchange rate determination that speaks to the predictability of forecast errors. The model follows Gourinchas and Tornell (2004), henceforth GT, in assuming that investors misperceive the relative importance of transitory and persistent interest rate shocks but departs from their setup in assuming that investors also over-extrapolate persistent interest rate shocks into the future. The model is kept simple to obtain analytical results. We leave the quantitative analysis to the general equilibrium (GE) model of Section 4 in which these belief distortions apply to the underlying economic fundamentals. The proofs of all Propositions are in Appendix B.

2.1 General Setup

Primitives. We begin with a log-linearized version of the standard foreign-exchange no-arbitrage condition:

$$\mathrm{E}_t^s\, s_{t+1} - s_t = x_t - \xi_t. \tag{8}$$

As in Section 1.2, E_t^s denotes the subjective expectations, which may differ from statistical or rational expectations. Equation (8) simply states that if the domestic interest rate differential exceeds the subjective expected depreciation of the domestic currency, then the domestic currency carries a positive *expected* excess return, ξ_t. Our GE model of Section 4 derives an analogous condition from first principles.

For now, we take the interest rate differential as the primitive of our model. The agents observe the realized interest rate differential, x_t, which is assumed to follow a first-order autoregressive process:

$$x_t = \rho x_{t-1} + \varepsilon_t, \qquad \varepsilon_t, \sim \mathcal{N}(0, \sigma_\varepsilon^2). \tag{9}$$

Equation (9) characterizes the true process. Agents observe the realization of x_t but, crucially, believe that the data-generating process is:

$$x_t = z_t + \nu_t, \qquad \nu_t \sim \mathcal{N}(0, \tilde{\sigma}_\nu^2), \tag{10}$$

$$z_t = \tilde{\rho} z_{t-1} + \varepsilon_t, \qquad \varepsilon_t, \sim \mathcal{N}(0, \sigma_\varepsilon^2). \tag{11}$$

Agents do not directly observe z_t. Instead, they infer it using the observation of x_t.

There are two differences between the true data generating process and the process perceived by the agents. First, the agents perceives the interest rate process to be buffeted also by a transitory shock, ν_t, with variance $\tilde{\sigma}_\nu^2$, instead of being a univariate AR(1) process. When $\tilde{\sigma}_\nu^2 > 0$, investors perceive interest rate changes to be more transitory than they actually are. This assumption is also present in GT and reflects the uncertainty surrounding the importance of different shocks that move interest rates.[14]

The second difference is that the persistent AR(1) component, z_t, is *perceived* to have an autocorrelation of $\tilde{\rho}$ instead of ρ. The case of $\tilde{\rho} > \rho$ captures over-extrapolation of today's interest rates to tomorrow, while $\tilde{\rho} < \rho$ encodes under-extrapolation. Over-extrapolative beliefs generate over-reaction to current disturbances and have been used in finance to explain a variety of asset-price phenomena (Barberis et al., 1998; Daniel et al., 1998; Hong and Stein, 1999; Greenwood and Shleifer, 2014; Gennaioli et al., 2016; Guo and Wachter, 2019). Conversely, under-extrapolative beliefs result in muted responses to shocks, are closer in spirit to level-K thinking or cognitive discounting (e.g., Gabaix, 2020), and are useful to explain the sluggish adjustment of macro variables to economic innovations.

We assume that agents act optimally conditional on their beliefs, but they do not update these beliefs over time to learn about the true model. We offer two motivations for this behavior. First, a departure from full rationality is necessary to explain the behavior of forecast errors documented in Section 1.2. If agents were rational, forecast errors would be unpredictable using observable variables such as the current interest rate differential, x_t. Second, these assumptions ensure minimal departure from a rational expectation model while allowing to parsimoniously capture several behavioral biases as we illustrate below.

Beliefs. Agents solve a standard signal-extraction problem to form beliefs about the path of future interest rates $\mathrm{E}_t^s\, x_{t+k}$. Conditional on these beliefs, the subjective no-arbitrage condition, (8), determines the equilibrium exchange rate. Given the linear-Gaussian environment described above, beliefs updating follows the standard Kalman Filter recursion derived, for example, in Hamilton (1994, Ch. 13).

Lemma 1. *Given an initial distribution of beliefs about* z_1 *that follows* $\mathcal{N}(\mathrm{E}_0^s\, z_1, \Sigma_1)$,

[14] We could have assumed that transitory shocks also affect the true data-generating process for x_t. We find this unnecessarily cumbersome as the algebra would become more complicated but our results would be qualitatively unchanged as long as the transitory component is perceived to be more volatile than it actually is. Furthermore, in Appendix D we show that the true interest rate data prefer an AR(1) representation to a representation of the form (10)-(11).

subjective expectations about the interest rate differential evolve according to:

$$\begin{aligned} \mathrm{E}^s_t\, x_{t+1} &= \tilde{\rho}\, \mathrm{E}^s_{t-1}\, x_t + \kappa_t \tilde{\rho}(x_t - \mathrm{E}^s_{t-1}\, x_t), \\ \kappa_t = \frac{\tilde{\rho}^2 \Sigma_t + \sigma_\varepsilon^2}{\tilde{\rho}^2 \Sigma_t + \sigma_\varepsilon^2 + \tilde{\sigma}_\nu^2}; &\quad \Sigma_{t+1} = (1-\kappa_t)(\tilde{\rho}^2 \Sigma_t + \sigma_\varepsilon^2). \end{aligned} \tag{12}$$

The Kalman gain κ_t and the conditional variance Σ_{t+1} eventually converge to their steady state values:

$$\kappa = \frac{\tilde{\rho}^2 \Sigma + \sigma_\varepsilon^2}{\tilde{\rho}^2 \Sigma + \sigma_\varepsilon^2 + \tilde{\sigma}_\nu^2}; \quad \Sigma = (1-\kappa)(\tilde{\rho}^2 \Sigma + \sigma_\varepsilon^2). \tag{13}$$

Equation (12) states that the forecast about future interest rate differential, $\mathrm{E}^s_t\, x_{t+1}$, is a weighted average of the forecast from the previous period, $\tilde{\rho}\, \mathrm{E}^s_{t-1}\, x_t$, and the current interest rate differential, x_t. The previous forecast is updated with the surprise in the observed interest rate relative to its forecast, $x_t - \tilde{\rho}\, \mathrm{E}^s_{t-1}\, x_t$. The Kalman gain κ_t represents the weight given to this surprise in updating the beliefs while Σ_{t+1} is the conditional variance of the persistent component of the interest rate differential, z_t.

The Kalman gain depends on the perceived relative importance of transitory and persistent shocks, $\frac{\tilde{\sigma}_\nu^2}{\sigma_\varepsilon^2}$, and on the perceived duration of the persistent component $\tilde{\rho}$. If transitory shocks are perceived to be relatively small (low $\frac{\tilde{\sigma}_\nu^2}{\sigma_\varepsilon^2}$) or shocks to z_t are thought to be very persistent (high $\tilde{\rho}$), a given innovation in x_t is more likely to have been caused by a shock to z_t, and $x_t - \mathrm{E}^s_{t-1}\, x_t$ would receive a larger weight κ_t.

Two important remarks are in order. First, the beliefs of the agents evolve according to the perceived model and, therefore, depend on the parameters $\tilde{\rho}$ and $\tilde{\sigma}_\nu^2$, rather than the parameters of the true data generating process. Second, as we show below, each belief distortion plays a distinct role on exchange rate determination, and is essential to explain the properties of forecast errors at different horizons.

We assume that at beginning of time agents are endowed with an infinite history of signals. This implies that the Kalman gain and the conditional variance have converged to their steady-state values.

Equilibrium exchange rate. By iterating forward the no-arbitrage condition (8) we obtain the solution for the exchange rate:

$$s_t = -\sum_{j=0}^{\infty} \mathrm{E}^s_t\, x_{t+j} + \sum_{j=0}^{\infty} \mathrm{E}^s_t\, \xi_{t+j} + \lim_{T\to\infty} \mathrm{E}^s_t\, s_{t+T}. \tag{14}$$

At this point, we make two simplifying assumptions that allow us to focus on misperceptions about the interest rate process. First, we abstract from sources of expected excess returns by setting $\xi_t = 0$, $\forall t$. Second, we assume that $\lim_{T\to\infty} \mathrm{E}^s_t\, s_{t+T} = 0$ so that the exchange rate is stationary and there are no misperceptions about its long-run level. These assumptions will be relaxed in the GE model in Section 4.

Proposition 1. *The equilibrium exchange rate is given by:*

$$s_t = -\sum_{j=0}^{\infty} \mathrm{E}^s_t\, x_{t+j} = -x_t - \frac{1}{1-\tilde{\rho}}\, \mathrm{E}^s_t\, x_{t+1}. \tag{15}$$

This equation reflects the asset view of exchange rates: s_t is determined by the cumulative sum of *perceived* interest rate differentials. Under rational expectations the agents misperceive neither the transitory component ($\tilde{\sigma}^2_\nu = 0$), nor the persistence of the long-lived component ($\tilde{\rho} = \rho$), so that $\mathrm{E}_t\, x_{t+1} = \rho x_t$, where E_t is the rational expectation operator. In this case, equation (15) simplifies to:

$$s^{RE}_t = -\frac{1}{1-\rho} x_t. \tag{16}$$

It is straightforward to see that under rational expectations there are no predictable excess returns at any horizons.

Lemma 2. *Under rational expectations, excess returns are unpredictable at all horizons:*

$$\beta_{k+1} \equiv \frac{\mathrm{Cov}\left(\Lambda_{t+k+1}, x_t\right)}{\mathrm{Var}(x_t)} = 0 \quad \forall k \geq 0.$$

Lemma 2 reflects a general property of rational expectations: future forecast errors cannot be predicted by current information.

Subtracting equation (15) from (16) we can write the equilibrium exchange rate as:

$$s_t = s^{RE}_t + \frac{1}{1-\rho}(\mathrm{E}_t\, x_{t+1} - \mathrm{E}^s_t\, x_{t+1}) - \frac{\tilde{\rho}-\rho}{(1-\tilde{\rho})(1-\rho)}\, \mathrm{E}^s_t\, x_{t+1}. \tag{17}$$

Equation (17) shows the two ways in which the equilibrium exchange rate deviates from the rational expectation exchange rate. First, whenever the interest rate differential increases, subjective interest rate forecasts partially under-react $\mathrm{E}_t\, x_{t+1} - \mathrm{E}^s_t\, x_{t+1} > 0$, because they partly attribute the interest rate increase to the subjective transitory component, ν_t. Other things equal, the nominal exchange rate appreciates less than under rational expectations. Second, because the long-lived component of the interest rate process is perceived to be more persistent than it actually is, subjective expectations over-react to the interest rate shock, as captured by the last component when $\tilde{\rho} > \rho$. Other things equal, this over-reaction makes the exchange rate appreciate more than under rational expectations following an interest rate shock. We now turn to explain why both under-reaction and over-reaction are necessary to explain the dynamics of forecast errors and excess returns.

2.2 Explaining the Puzzles

We begin by recursively substituting the belief updating equation (12) into the expression for the nominal exchange rate (15) to obtain:

$$s_t = -x_t - \frac{\tilde{\rho}}{1-\rho}\kappa \sum_{i=0}^{\infty} [(1-\kappa)\tilde{\rho}]^i x_{t-i}.$$

If the interest rate differential follows an AR(1) process we can obtain the following analytical expression for the coefficients of the excess returns regressions.

Proposition 2. *Under shock misperception and over-extrapolation, the coefficients of the excess returns predictability regression (i.e., equation (3)) take the form:*

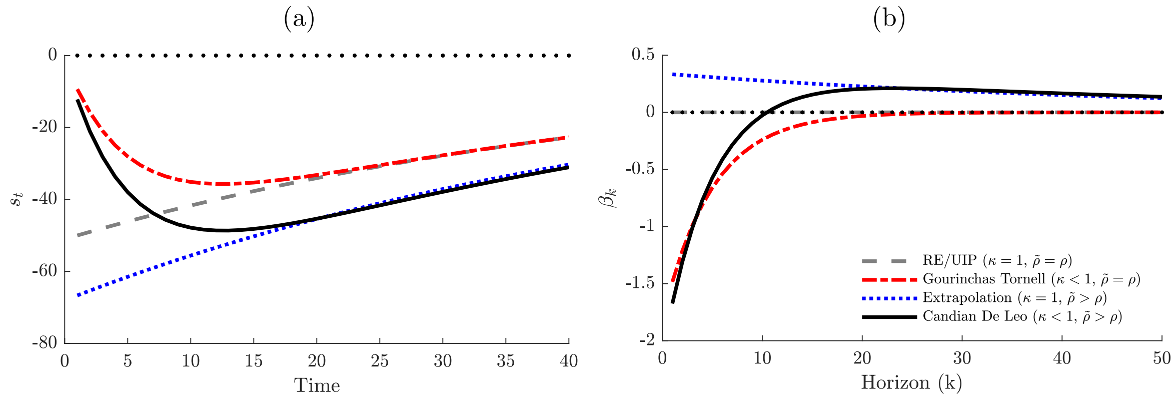
$$\beta_{k+1} = \rho^k \frac{1-(1-\kappa)\tilde{\rho}}{(1-\tilde{\rho})\left(1-(1-\kappa)\frac{\tilde{\rho}}{\rho}\right)} \left\{ (\tilde{\rho}-\rho) + \tilde{\rho}\kappa \left[(1-\kappa)\frac{\tilde{\rho}}{\rho} \right]^k \frac{(1-\kappa)\tilde{\rho}(\rho-\rho^{-1})}{1-(1-\kappa)\tilde{\rho}\rho} \right\}. \quad (18)$$

With this expression at hand we are now ready to explain the exchange rate puzzles. We begin by describing the delayed overshooting of the exchange rate to an interest rate shock, which is a useful stepping stone to understand the other puzzles. In describing the intuition of the results we always consider an increase in the relative domestic interest rate x_t, which leads to a domestic appreciation (a fall in s_t).

Delayed Overshooting and Magnified Adjustment The notion of delayed overshooting was first introduced by Eichenbaum and Evans (1995), who documented that following an unanticipated domestic monetary tightening, the nominal exchange rate appreciates for several periods before gradually depreciating.

Figure 2a shows the response of the exchange rate to a domestic interest rate shock in our model (black line) for the case with $\rho = 0.98$, $\tilde{\rho} = 0.985$ and $\kappa = 0.17$, along

Figure 2: UIP Regression Results (1990:M1-2007:M12)



Notes: Panel (2a) depicts the impulse response function of the exchange rate to a unitary increase in the domestic interest rate differential. Panel (2b) depicts the β_k coefficients from the excess return predictability regression (3). The parameters are set as follows: $\rho = 0.980$, $\tilde{\rho} = 0.985$, and $\kappa = 0.167$.

with the responses under special cases of our model. The values of ρ and κ are among those considered in GT, while the value of $\tilde{\rho} > \rho$ reflects our assumption of over-extrapolation. The figure is purely illustrative as our results in this Section are fully analytical and derived in the various propositions that follow.

If expectations were rational (dashed line in Figure 2a), the nominal exchange rate response would simply be a scaled-up version of this AR(1) interest rate response as per equation (16). That is, the exchange rate would immediately appreciate beyond its long run equilibrium value and then gradually depreciate toward this value, exhibiting the standard Dornbusch (1976) overshooting behavior.

With distorted beliefs, $\kappa < 1$ and $\tilde{\rho} > \rho$, and the response of the exchange rate to a monetary innovation ε_t is given by:

$$s_{t+j} = -\rho^j \left\{ 1 + \frac{\tilde{\rho}}{1-\tilde{\rho}} \left(\kappa \frac{1 - \left[(1-\kappa)\frac{\tilde{\rho}}{\rho} \right]^{j+1}}{1 - (1-\kappa)\frac{\tilde{\rho}}{\rho}} \right) \right\} \varepsilon_t. \tag{19}$$

Initially, investors do not know whether the interest rate increase is due to a temporary shock or a persistent one, and they believe that the persistent shock has an autocorrelation of $\tilde{\rho}$. Two contrasting forces affect the initial exchange rate response. On the one hand, uncertainty about whether the shock is transitory or persistent (i.e., $\kappa < 1$) leads to a smaller appreciation than under rational expectations. On the other hand, conditional on the shock being a persistent one, agents tend to overestimate its persistence as $\tilde{\rho} > \rho$. Ceteris paribus, this second force leads to a larger appreciation than under rational expectations. Under any reasonable parameterization the first force dominates in the initial periods after the shock, leading to a muted appreciation.

Owing to the persistence of the interest rate shock, for several periods agents observe a domestic interest rate that exceeds their prior beliefs. The consequent revision of their beliefs leads to further appreciations of the exchange rate along the impulse response, as in GT. Differently from GT, when agents eventually learn that the shock was truly persistent, they attribute too much persistence to it as $\tilde{\rho} > \rho$. Thus, the exchange rate exhibits a magnified adjustment and becomes more appreciated than under rational expectations before reverting back to its steady state.

Equation (19) shows that it is the uncertainty about the type of shock that occurred that causes the delayed overshooting. Indeed, if agents knew for certain that the shock was persistent ($\kappa = 1$), the exchange would overshoot immediately at the time of the shock. The over-extrapolation would simply scale up the rational expectation response, since the undiscounted sum of future interest rates is believed to be larger than in reality (blue line). By contrast, in a GT economy with only misperception about the type of shock (red line), the initial response of the exchange rate would be weaker than under rational expectation and it would gradually catch up with the rational expectation response, yet without ever exceeding it. Next, we will see that both the initial under-reaction and subsequent over-reaction of the exchange rate relative to the rational expectation response, and thus both behavioral biases, are needed to explain the empirical puzzles of Section 1.

Fama Puzzle Setting $k = 1$ in Proposition 2 we obtain the coefficient for the classical Fama regression (2) given in the following proposition.

Proposition 3. *The Fama predictability coefficient in the model is:*

$$\beta_1 = \frac{1-(1-\kappa)\tilde{\rho}}{1-\tilde{\rho}}\left(\frac{\tilde{\rho}\kappa}{1-(1-\kappa)\tilde{\rho}\rho} - \rho\right).$$

The Fama coefficient is negative as long as $\tilde{\rho} < \bar{\rho} \equiv \left[(1-\kappa)\rho + \frac{\kappa}{\rho}\right]^{-1}$. *This condition is likely to be satisfied if* $\tilde{\rho}$ *is not much larger than* ρ.

Intuitively, the Fama coefficient is negative as long as there is a sufficient degree of under-reaction. Under-reaction, which dominates in the short run, leads agents to revise the path of interest rate differentials upward for some periods following an interest rate shock. Such revisions cause the exchange rate to appreciate in the short run while the interest rate is high (Figure 2a). These patterns imply an unconditional negative correlation between interest rate differentials and next-period excess returns ($\beta_1 < 0$). The initial delayed overshooting of the exchange rate is thus responsible for $\beta_1 < 0$, which arises when over-extrapolation is not too powerful.[15]

Predictability Reversal The properties of excess returns predictability at longer horizons, measured by β_k, are summarized in the following proposition.

Proposition 4. *Define* $\bar{\rho} \equiv \left[(1-\kappa)\rho + \frac{\kappa}{\rho}\right]^{-1}$. *The following holds true for* β_k *as defined in Lemma 2:*

- *If* $\tilde{\rho} \leq \rho$, *then* β_k *is negative for all* $k \geq 1$.
- *If* $\rho < \tilde{\rho} < \bar{\rho}$, *then* $\exists$ *a* $\bar{k} > 1$ *such that* β_k *is negative for* $k < \bar{k}$ *and positive for* $k \geq \bar{k}$. β_k *converges to zero as* $k \to \infty$.
- *If* $\tilde{\rho} \geq \bar{\rho}$, *then* β_k *is positive for all* $k \geq 1$.

Proposition 4 conveys two important results. First, it shows that both behavioral biases are necessary to deliver the predectability reversal.[16] Figure 2b illustrates this result: only when both biases are active does the profile of β_k cross the zero. Second, in the presence of over-extrapolation ($\tilde{\rho} > \rho$), if the model explains the classic UIP puzzle (which requires $\tilde{\rho} < \bar{\rho}$ as per Proposition 3), it will also be able to explain the predictability reversal puzzle. That is, in response to an increase in the domestic interest rate, the domestic currency initially earns positive excess returns, while these returns subsequently turn negative. These patterns of excess returns obtain because of the relative importance of the forces that operate at different horizons. Initially, Bayesian learning dominates and thus agents underestimate the future path of the interest rate differential. Later, the dominant force is over-extrapolation, which results in agents over-estimating the future path of the interest rate differential.

Which force dominates at a given horizon can be seen analytically from the expression of β_k in (18). Note there that the term in curly brackets is increasing in

[15] Granziera and Sihvonen (2020) and Valente et al. (2021) highlight that under-reaction of short-term interest rate beliefs can also explain the downward-sloping term structure of UIP violations (Lustig et al., 2019).

[16] In fact, when $\kappa = 1$, then $\bar{\rho} = \rho$ and the conditions for the reversal are thus met.

k. Recalling that under rational expectations $\beta_k = 0$ for all k, it is easy to see that under-reaction dominates while the bracketed terms are negative, while over-reaction does when those terms are positive.

Figure 2b illustrates with an example that models with only misperception about transitory shocks or with only over-extrapolation cannot explain the entire profile of $\beta_k' s$. Appendix C proves that this is generally the case for these models as well as for models with simple diagnostic expectations as in Bordalo et al. (2018).

Excess Comovement Puzzle The excess comovement puzzle says that high interest rate currencies tend to be strong relative to the UIP exchange rate. Formally, this can be expressed as:

$$\text{Cov}(s_t - s_t^{IP}, x_t) < 0.$$

Engel (2016) shows that this inequality holds in the data for 6 currencies.[17] Equivalently, this condition can be stated as:

$$\sum_{k=1}^{\infty} \beta_k > 0. \tag{20}$$

Equation (20) reveals that if $\beta_1 < 0$, then the predictability reversal (i.e., $\beta_k > 0$ at some $k > 1$) is a necessary (yet not sufficient) condition to generate the excess comovement (Engel, 2016). The next Proposition shows that our model can deliver excess comovement.

Proposition 5. *The model with shock misperception and over-extrapolation explains the excess comovement puzzle (i.e., it satisfies equation* (20)*) as long as:*

$$\tilde{\rho} - \rho > \frac{\tilde{\rho}\kappa(1-\rho)}{1-(1-\kappa)\tilde{\rho}} \left(\frac{(1-\kappa)\tilde{\rho}(\rho^{-1}-\rho)}{1-(1-\kappa)\tilde{\rho}\rho} \right) > 0. \tag{21}$$

To understand Proposition 5, recall from Proposition 4 that whenever our model explains the predictability reversal, the β_k's switch sign once and for all after horizon $\bar{k}$. It follows that equation (20) holds whenever the positive excess returns (for $k > \bar{k}$) more than offset the negative ones (for $k < \bar{k}$). This obtains when the phase of over-reaction more than compensates for the initial under-reaction. The condition (21) shows formally that over-extrapolation is necessary to explain the excess comovement

[17] Engel (2016) works with the real version of this inequality.

puzzle because it is satisfied only if $\tilde{\rho} > \rho$. At the same time, over-extrapolation must not be excessively dominant in order for the model to feature enough initial under-reaction to explain the Fama puzzle (Proposition 3). Overall then, both shock misperception over-extrapolation must be operative to account for the broad set of exchange rate puzzles we have examined. For this reason, models with only one bias fall short of explaining at least one of these defining properties of exchange rate dynamics (See Figure 2b and Appendix C).

3 Testing the Mechanism

In this section, we provide empirical support for our proposed model of expectations formation by testing one of its key predictions for interest rates forecast errors. In contrast to exchange rate expectations, interest rates expectations were not a target of our analysis so far and thus constitute a validation test for our theory. Our model predicts that, following an innovation in the interest rate differential, interest rate expectations initially under-react, while subsequently over-react. We now show that this property of interest rate expectations is necessary to obtain the reversal in exchange rates forecast errors and that it actually holds in the data.

We first characterize the properties of interest rate forecast errors in our model.

Proposition 6. *Define χ_k as the Impulse Response Function (IRF) of the one-step ahead forecast error of the interest rate differential to a time-t innovation in the interest rate differential (Eq. 9). That is:*

$$\chi_k = \frac{\partial(x_{t+k} - \mathrm{E}^s_{t+k-1}\, x_{t+k})}{\partial \varepsilon_t} \quad \forall k \geq 0.$$

Define $\hat{\rho} \equiv \rho/\kappa$. The following holds true for χ_k in the model of Section 2:

- *If $\tilde{\rho} \leq \rho$, then $\chi_k > 0$ for all $k \geq 1$.*
- *If $\rho < \tilde{\rho} < \hat{\rho}$, then $\exists$ a $\hat{k} > 1$ such that $\chi_k > 0$ for $k < \hat{k}$ and $\chi_k < 0$ for $k > \hat{k}$.*
- *If $\tilde{\rho} \geq \hat{\rho}$, then $\chi_k < 0$ for all $k \geq 1$.*

Proposition 6 demonstrates that both misperception about transitory shocks ($\kappa < 1$) and over-extrapolation ($\tilde{\rho} > \rho$) are required to get a sign switch in the forecast errors for interest rates. Indeed, without over-extrapolation the IRF of the interest rate forecast error is always positive, while if there is over-extrapolation but no shock misperception the IRF is always negative.[18] Only when both are present will the impulse response switch sign.

The previous discussion indicates that the behavior of the IRF in the data is informative about the underlying frictions driving interest rates forecast errors. Accordingly, we estimate IRFs of interest rate differential forecast errors to an innovation in the empirical process for the interest rate differential. To construct interest rates forecast errors, we use survey data on 3-month-ahead interest rate expectations from *FX4Casts* from 1986:M8 to 2007:M12 for Canada, Germany, Japan, Switzerland and United Kingdom as well as the United States. These are the countries for which we have a sufficient amount of observations about interest rate expectations.

We estimate impulse responses using the projection method of Jordà (2005). The estimating equation, for each horizon $1 < k < K$, is:

$$x_{j,t+k} - \mathrm{E}^{s}_{t+k-1}\, x_{j,t+k} = \psi_{j,k} + \chi_{j,k}\hat{\varepsilon}_{j,t} + u_{j,t+k}, \tag{22}$$

where $\hat{\varepsilon}_{j,t}$ is the estimated innovation in the interest rate process when we fit (10)-(11) to the actual interest rate data.[19] The sequence $\{\chi_{j,k}\}$ is an estimate of the IRF of the interest rate differential forecast errors to an innovation in the interest rate differential.

Figure 3 reports the impulse response estimates. We report the panel estimates as well as their 95 percent confidence intervals, along with the country-by-country estimates.[20] The panel coefficients are positive and statistically significant at horizons of up to around one year, but they change sign at longer horizons, becoming negative and statistically significant at horizons between 3 to 5 years. These estimates imply that following a positive innovation in the interest rate differential innovation agents' interest rate differential expectations initially under-react, while subsequently over-react.[21] This evidence, which emerges also in the country-level regressions, supports the presence of both shock misperception and over-extrapolation in the data.[22]

[18]Without beliefs distortion $\kappa = 1 \implies \hat{\rho} = \rho$ so that the third case of Proposition 6 applies.

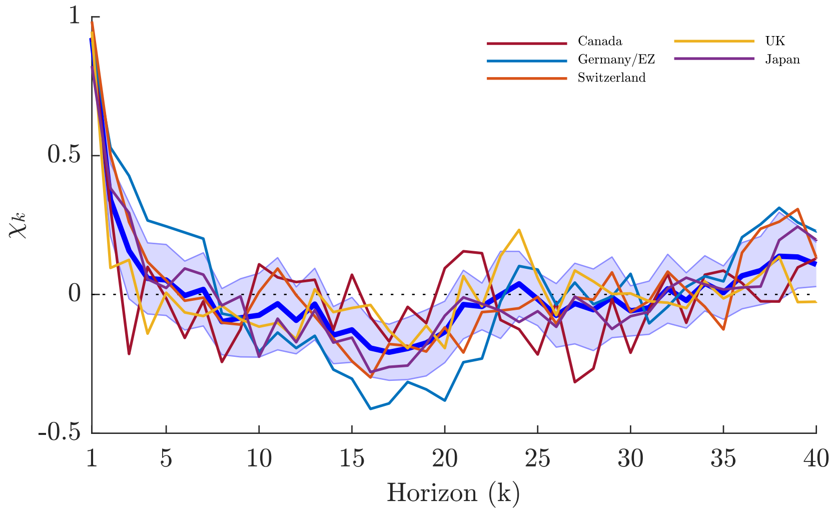
[19]The estimates of (10)-(11) on the actual interest rate data support the hypothesis that $\sigma^2_{\nu} = 0$, thus implying that the interest rate differential follows an AR(1) process with unique innovation ε_t. See Appendix D for more details.

[20]The standard errors are computed following Driscoll and Kraay (1998) to correct for heteroskedasticity, and both serial and cross-equation correlation.

[21]Figure A.2 shows similar results for the sample period that include the zero lower bound.

[22]Our evidence does not rely on the identification of a particular shock, as the theoretical predictions of Proposition 6 and Corollary 1 hold unconditionally. Similar evidence holds for the dynamic response of interest rate forecast errors to identified monetary policy shocks (Valente et al., 2021).

Figure 3: IRFs of Forecast Errors in Interest Rate Differentials (1986:M8-2007:M12)



Notes: The figure reports the empirical impulse response of the one-quarter ahead forecast error in the interest rate differential to an innovation in the interest rate differential (see Eq. (22)). Each horizon is a quarter.

How does the IRF of the interest rates forecast error relates to the reversal of the coefficients β_k from the excess return predictability regression that motivated our model? We address this question in the following Corollary.

Corollary 1. *The coefficients β_k in the excess return predictability regression switch sign from negative to positive only if the coefficients χ_k switch sign from positive to negative.*

It should come at no surprise that a sign switch in the IRF for interest rate forecast errors is necessary for a sign switch in the excess return predictability coefficients, as both phenomena require the same two departures from rational expectations. Therefore, the findings in Figure 3 directly supports the explanation we have provided in this paper for the excess return predictability reversal. The complementary lesson we draw in this section is that the mere presence of these two ingredients is not sufficient to obtain reversal of excess returns predictability; they should also balance themselves out in a specific manner. In particular, over-extrapolation should not have too strong of a bite at short horizons but should dominate only later on.

4 Distorted Beliefs in General Equilibrium

We now embed our beliefs distortions into a general-equilibrium (GE) environment where interest rates and expectations thereof are determined endogenously. The analysis serves several purposes. First, it demonstrates that the mechanism we proposed to address UIP deviations at different horizons survives GE effects, and that the reversal of excess currency returns predictability is a pattern that arises under both supply and demand shocks typically considered in open-economy models. Second, the model reproduces several exchange-rate related moments of interest in open-economy macroeconomics. Third, it shows that the belief distortions proposed here are consistent with the recently documented properties of forecast errors for several macro variables such as output, inflation, and interest rates (Coibion and Gorodnichenko, 2015; Angeletos et al., 2020; Kohlhas and Walther, 2021) as well as recent evidence from survey data on exchange rate forecasts (Stavrakeva and Tang, 2020b), providing further support to our proposed expectations formation process. Fourth, it proposes a model in which belief distortion coexist with time-varying liquidity yields in driving predictable excess returns, and replicating the different aspects of the Froot and Frankel (1989) regression. Fifth, it shows that excess currency returns in our GE model are endogenous to the policy regime in a way that jointly rationalizes the evidence about the direction of UIP deviations before *and* after the global financial crisis (Bussiere et al., 2018).

4.1 A New-Keynesian Open-Economy Model

The framework is a standard New-Keynesian open-economy model that features two countries of equal size populated by households, a continuum of monopolistically competitive producers, and a monetary authority. Each country specializes in the production of one type of tradable goods, produced in a number of varieties. All goods are traded and consumed in both countries, with home bias in consumption preferences. Prices are sticky and set in the producer's currency. Households earn a convenience yield on government bonds, which gives rise to endogenous time-varying expected excess currency returns. Convenience yields are not necessary for our main results, but allow us to have a model counterpart for the ex-ante excess currency returns that we see in the data. Monetary policy follows a conventional Taylor rule targeting inflation, resulting in a floating nominal exchange rate.

The economies are buffeted by TFP and demand shocks. The main departure from standard models lies in the shocks expectations formation process, which follows the same characteristics of the partial equilibrium model of Section 2. We now present the structure of the home economy in more detail. Unless otherwise specified, the foreign economy is symmetric, and foreign variables will be denoted with an asterisk.

Households The utility function of the representative household in country H is

$$\mathrm{E}_0^s \sum_{t=0}^{\infty} (\beta^t \zeta_t) \left[\frac{C_t^{1-\sigma} - 1}{1-\sigma} - \frac{L_t^{1+\varphi}}{1+\varphi} + \frac{D_t^{1-\iota}}{1-\iota} \right], \tag{23}$$

where

$$D_t = \frac{M_{H,t+1},}{P_t} + \eta \frac{B_{H,t+1}}{P_t}. \tag{24}$$

Above, $\beta < 1$ is the discount factor and E_t^s is the subjective expectation operator described more formally below. Households receive utility from consumption, C_t, and disutility from working, where L_t indicates hours of labor input in the production of domestic varieties. The parameter $\varphi > 0$ is the inverse of the Frisch elasticity of labor supply and σ denotes the inverse of the intertemporal elasticity of substitution. The preference or "demand" shock, ζ_t, makes the domestic household value present-period utility more than future utility.

The world economy features six assets: home country money (M_t), foreign country money (M_t^*), home government bonds (B_t), foreign government bonds (B_t^*), home country "market" bonds (B_t^m), and foreign country "market" bonds (B_t^{m*}).

Following Nagel (2016) and Engel (2016), home households derive liquidity services, D_t, from both home money and home government bonds, as shown in equation (24). We assume $0 \leq \eta < 1$, so that B_t is less liquid than M_t. We will assume that the supply of the assets and the parameterization of the utility function are such that the home household always holds home money and home government bonds. The home household cannot trade foreign government bonds or foreign money.

Households consume both domestically produced and imported goods, respectively, $C_t(h)$ and $C_t(f)$. We assume that each good h (or f) is an an imperfect substitute for all other goods' varieties, with constant elasticity of substitution $\nu > 1$:

$$C_{Ht} \equiv \left(\int_0^1 C_t(h)^{\frac{\nu-1}{\nu}} dh \right)^{\frac{\nu}{\nu-1}}, \quad C_{Ft} \equiv \left(\int_0^1 C_t(f)^{\frac{\nu-1}{\nu}} df \right)^{\frac{\nu}{\nu-1}}.$$

The overall consumption baskets, C_t, pools home and foreign goods according to:

$$C_t \equiv \left((1-\gamma)^{\frac{1}{\theta}}(C_{Ht})^{\frac{\theta-1}{\theta}} + \gamma^{\frac{1}{\theta}}(C_{Ft})^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}} \qquad \theta > 0,$$

where $\gamma \in [0, 1/2)$ governs the home bias in consumption preferences, and θ is the elasticity of substitution between home and foreign goods, or trade elasticity. The utility-based consumption price index is $P_t = \left((1-\gamma)P_{Ht}^{1-\theta} + \gamma P_{Ft}^{1-\theta}\right)^{\frac{1}{1-\theta}}$, where $P_{Ht} = \left(\int_0^1 p_t(h)^{1-\nu} dh\right)^{\frac{1}{1-\nu}}$ and $P_{Ft} = \left(\int_0^1 p_t(f)^{1-\nu} df\right)^{\frac{1}{1-\nu}}$ are the domestic-currency price indices for home and foreign goods.

Let S_t denote the nominal exchange rate expressed in domestic currency per foreign currency, so that an increase in S_t represents a depreciation of the home currency. Correspondingly, the real exchange rate is $\mathcal{Q}_t \equiv \frac{S_t P_t^*}{P_t}$ and an increase in $\mathcal{Q}_t$ indicates a real domestic depreciation. Because prices are set in the currency of the producer, the law of one price holds so that $P_t(h) = S_t P_t(h)^*$ and $P_{Ht} = S_t P_{Ht}^*$. Nevertheless, home bias in consumption will result in deviations of the real exchange rate from purchasing-power parity, i.e., $\mathcal{Q}_t \neq 1$. Using the law of one price, we can link the the real exchange rate to the terms of trade, $\mathcal{T}_t \equiv \frac{P_{Ft}}{S_t P_{Ht}^*}$, which in log linear form is $q_t = (1-2\gamma)t_t$. Throughout the paper, lower-case letters denote percentage deviations from steady state, assuming symmetric initial conditions.

Total demand for a generic home and foreign variety can be written as:

$$Y_t^d(h) = \left(\frac{P_t(h)}{P_{Ht}}\right)^{-\nu} \left(\frac{P_{Ht}}{P_t}\right)^{-\theta} \left[(1-\gamma)C_t + \gamma \mathcal{Q}_t^{\theta} C_t^*\right], \tag{25}$$

$$Y_t^d(f) = \left(\frac{P_t(f)^*}{P_{Ft}^*}\right)^{-\nu} \left(\frac{P_{Ft}^*}{P_t^*}\right)^{-\theta} \left[\gamma \mathcal{Q}_t^{-\theta} C_t + (1-\gamma)C_t^*\right].$$

Budget constraint The home household uses its revenues in every period to purchase consumption goods or invest in home country money (M_t), home government bonds (B_t), home "market" bonds (B_t^m), and foreign "market" bonds (B_t^{m*}). Following Turnovsky (1985) and Benigno (2009), we assume that households incur a quadratic cost in changing the real asset position with respect to a constant real value (b^{m*}) when trading market bonds from abroad. This assumption guarantees the stationarity of net foreign asset positions.[23] We denote the gross nominal interest rate on home

[23]For simplicity, we assume that the adjustment costs paid by domestic households and foreign households accrue as profits to financial intermediaries and are entirely rebated to foreign households.

government bonds with R_t (and, thus, $1/R_t$ is the price of the bond), while the gross nominal interest rate on home and foreign country "market" bonds with R_t^m and R_t^{m*}, respectively. The home household's budget constraint is:

$$
\begin{aligned}
&C_t + \frac{M_{H,t+1}}{P_t} + \frac{B_{H,t+1}}{R_t P_t} + \frac{B^m_{H,t+1}}{R^m_t P_t} + \frac{S_t B^{m*}_{H,t+1}}{R^{m*}_t P_t} + \frac{\delta}{2}\left(\frac{S_t B^{m*}_{H,t+1}}{P_t} - b^{m*}\right)^2 \leq \\
&\frac{W_t}{P_t} L_t + \frac{M_{H,t}}{P_t} + \frac{B_{H,t}}{P_t} + \frac{B^m_{H,t}}{P_t} + \frac{S_t B^{m*}_{H,t}}{P_t} + \Pi_t + T_t^M - T_t.
\end{aligned}
\tag{26}
$$

Here, W_t is the wage, Π_t are real profits of domestic firms, and T_t are real government lump-sum taxes. Home (foreign) holding of assets are denoted with a subscript H (F). T_t^M denotes seignorage revenues from printing domestic money, which for simplicity are rebated in a lump-sum fashion to the household. Households maximize (23) subject to (26), and to the constraints $M_{H,t+1} \geq 0$, $B_{H,t+1} \geq 0$. These latter constraints mean that households are unable to issue securities with the same liquidity properties as government securities. The households' optimality conditions are in Appendix E.1.

Firms Domestic producers sell differentiated goods under monopolistic competition, facing the demand function (25). Their production function is the following:

$$
Y_t(h) = A_t L_t(h), \tag{27}
$$

where $L_t(h)$ denotes labor services employed by firm h in period t, and A_t represents aggregate domestic total factor productivity. We introduce price stickiness in a conventional way and denote with λ_p the Calvo probability of price non-adjustment. When firm h has the opportunity, it sets the domestic-currency price $\tilde{P}_t(h)$ to maximize the expected discounted value of net profits:

$$
\max_{\tilde{P}_t(h)} \mathrm{E}_t^s \sum_{s=0}^{\infty} (\lambda_p)^s \Theta_{t,t+s} \left[\tilde{P}_t(h) Y^d_{t+s}(h) - W_{t+s} L_{t+s}(h)\right],
$$

subject to the demand function (25) and the production function (27). Profits are discounted using the home households' stochastic discount factor: $\Theta_{t,t+s} = \beta \frac{\zeta_{t+s}}{\zeta_t} \left(\frac{C_{t+s}}{C_t}\right)^{-\sigma} \frac{P_t}{P_{t+s}}$.

Fiscal and monetary policy The government is assumed to finance its debt by means of lump-sum taxes T_t. The governenment budget constraint is given by:

$$\frac{B_{H,t+1}}{R_t} + T_t = B_{H,t}. \tag{28}$$

Monetary policy is conducted by adjusting the short-term nominal rate on government bonds according to the following rule:

$$r_t = \phi_\pi \pi_t,$$

where $\pi_t = \Delta \log P_t$ is the (log of) domestic CPI inflation and $r_t = \log R_t$.

Shocks and expectations The domestic and foreign economies are buffeted by preference and productivity shocks. Each of these exogenous disturbances, which we denote with $x_t \in \{\log A_t, \log A_t^*, \log \zeta_t, \log \zeta_t^*\}$, follows an AR(1) process in logs with persistence parameter ρ_x and variance of innovations $\sigma_x^2 \geq 0$:

$$x_t = \rho_x x_{t-1} + \varepsilon_t^x, \qquad \varepsilon_t^x, \sim \mathcal{N}(0, \sigma_x^2). \tag{29}$$

where the innovations ε_t^x are assumed to be *i.i.d.* and uncorrelated across countries. Households and firms observe the realizations of the shocks but have the following beliefs about the exogenous process in (29):

$$x_t = z_t^x + \nu_t^x, \qquad \nu_t^x \sim \mathcal{N}(0, \tilde{\sigma}_{\nu,x}^2), \tag{30}$$

$$z_t^x = \tilde{\rho}_x z_{t-1}^x + \varepsilon_t^x, \qquad \varepsilon_t^x, \sim \mathcal{N}(0, \sigma_x^2). \tag{31}$$

The beliefs can be indifferently parameterized by the couple $(\tilde{\rho}_x, \tilde{\sigma}_{\nu,x}^2/\sigma_x^2)$ or $(\tilde{\rho}_x, \kappa_x)$, where κ_x denotes the Kalman gain associated with the signal extraction problem in (30)-(31). We choose to work with the latter parameterization for convenience.

Market clearing, equilibrium, and solution method In equilibrium, households and firms optimally choose quantities and prices subject to their respective constraints while markets clear. Labor market clearing requires that L_t equals simultaneously the labor supply of the households and the total labor demand of the firms defined as $\int_0^1 L_t(h)dh$, and equivalently for L_t^* in the foreign economy. The wage rates, W_t and W_t^*, adjust to ensure that these markets clear. Good markets clear, ensuring that production equals total demand for domestic and foreign variety, i.e., $Y_t(h) = Y_t^d(h)$ and $Y_t(f) = Y_t^d(f)$. Defining the indexes for domestic and foreign produced goods as

$Y_{Ht} = \left(\int_0^1 Y_t(h)^{\frac{\nu-1}{\nu}} dh\right)^{\frac{\nu}{\nu-1}}$ and $Y_{Ft} = \left(\int_0^1 Y_t(f)^{\frac{\nu-1}{\nu}} df\right)^{\frac{\nu}{\nu-1}}$ we obtain:

$$Y_{Ht} = \left(\frac{P_{Ht}}{P_t}\right)^{-\theta} \left[(1-\gamma)C_t + \gamma \mathcal{Q}_t^{\theta} C_t^*\right], \quad Y_{Ft} = \left(\frac{P_{Ft}^*}{P_t^*}\right)^{-\theta} \left[\gamma \mathcal{Q}_t^{-\theta} C_t + (1-\gamma)C_t^*\right].$$

Finally, the markets for all the assets, B_t, B_t^*, M_t, M_t^*, B_t^m and B_t^{m*}, have to clear in equilibrium. We solve the model using a log-linear approximation of the equilibrium conditions around a deterministic and symmetric zero-inflation steady state. The solution method, which accommodates distorted beliefs, is described in Appendix E.5.

Uncovered interest rate parity In Appendix E, we show that up to a first-order approximation the households' optimality imply the following modified UIP condition:

$$\mathrm{E}_t^s \, \Delta s_{t+1} = (r_t - r_t^*) + \alpha(r_t - r_t^*) + \delta b_{t+1}^m, \tag{32}$$

where $\alpha = \beta \frac{\eta}{1-\eta} > 0$. The no-arbitrage condition assumed in our partial equilibrium analysis of Section 2 is a special case of equation (32) when there are no intermediation costs and government bonds do not provide liquidity services, i.e., $\delta = \alpha = 0$. Since households derive liquidity services from government bonds, equation (32) features time-varying *expected* excess returns. When home bonds provide larger liquidity services, the foreign bonds must pay a higher expected monetary return. As in Engel (2016) and Engel and Wu (2020), relative liquidity services are proportional to the interest rate differential. Because home money and home bonds are imperfect substitutes in providing liquidity services to the home household (see eq. (24)), a higher home interest rate raises the opportunity cost of holding home money, making home government bonds become more valued for their liquidity services.

Calibration Our calibration follows a two-pronged approach. First, we calibrate the standard parameters of open-economy models to conventional values following the broader macro literature. Second, we discipline the extent of belief distortions so that the model reproduces some empirical facts about the predictability of forecast errors of U.S. output and U.S. inflation recently documented in the literature.

The calibration of standard parameters is summarized in Panel A of Table 1. We set the Frisch elasticity of labor supply $1/\varphi = 1$, the quarterly discount factor $\beta = 0.99$, and the inverse of the intertemporal elasticity of substitution $\sigma = 5$. We fix the home

Table 1: Baseline Calibration

Parameter	Interpretation	Value	Parameter	Interpretation	Value
A. Preferences and Technology			ϕ_π	Monetary policy inflation response	2.50
β	Discount factor	0.99	α	Elasticity of liquidity function	0.15
σ	Risk aversion	5.00			
φ	Inverse Frisch elasticity	1.00	*B. Shocks and Beliefs*		
λ_p	1- Probability of price reset	0.75	σ_ζ/σ_a	Relative std of shocks	2.25
γ	Openness	0.05	ρ_x	Shocks persistence	0.90
θ	Trade elasticity	1.50	$\tilde{\rho}_x$	Shocks perceived persistence	0.98
δ	Intermediation cost	0.001	κ_x	Kalman gain	0.47

bias parameter $\gamma = 0.05$. For the trade elasticity θ, we choose the value of 1.5, as originally chosen by Backus et al. (1994) and Chari et al. (2002). We set the Taylor-rule parameter ϕ_π to 2.5. The Calvo probability of non-price-reset is $\lambda_p = 0.75$. We follow Engel (2016) in setting $\alpha = 0.15$.

Regarding the exogenous processes, we assume that both productivity and preference shocks have persistence of 0.90, and we calibrate the relative standard deviation of preference shocks to productivity shocks in order to match the well-documented empirical negative correlation between consumption differentials and real exchange rates: $corr(\Delta c_t - \Delta c_t^*, \Delta q_t) = -0.17$. Absent more guidance from the data, we restrict the belief distorsions (κ_x and $\tilde{\rho}_x$) to be the same for both sets of shocks. We then choose values of κ_x and $\tilde{\rho}_x$ so that the model reproduces some well-known empirical facts about the predictability of forecast errors of U.S. output and U.S. inflation. In particular, consider the following two regression equations:

$$x_{t+3} - \mathrm{E}_t^s\, x_{t+3} = \alpha^{CG} + \beta^{CG}(\mathrm{E}_t^s\, x_{t+3} - \mathrm{E}_{t-1}^s\, x_{t+3}) + \varepsilon_{t+3}^{CG}; \tag{33}$$

$$x_{t+3} - \mathrm{E}_t^s\, x_{t+3} = \alpha^{KW} + \beta^{KW} x_t + \varepsilon_{t+3}^{KW}. \tag{34}$$

Equation (33) consists in a regression of forecast errors about variable x on its forecast revision. Full-information rational expectations imply $\beta^{CG} = 0$, while the literature has documented that β^{CG} is positive and significant for several macro variables, including inflation, output growth and unemployment forecasts (Coibion and Gorodnichenko, 2015; Angeletos et al., 2020; Kohlhas and Walther, 2021).[24] That is, when professional forecasters, in aggregate, revise upward their estimation of output or inflation, they on average always "undershoot" the eventual truth.

Equation (34) consists in a regression of forecast errors about variable x on its

[24] As Coibion and Gorodnichenko (2015), we specify the regression with 3-quarters-ahead forecasts.

Table 2: Forecast-error Moments

	Data	Model		Data	Model
A. Coibion-Gorodnichenko Regressions			*B. Kohlhas-Walther Regressions*		
β^{CG} Inflation	0.60 (0.53)	0.80 (0.92)	β^{KW} Inflation	-0.13 (0.08)	-0.03 (0.08)
β^{CG} Output	0.66 (0.20)	0.66 (0.24)	β^{KW} Output	-0.08 (0.05)	-0.13 (0.07)
β^{CG} Interest rate	0.64 (0.17)	0.80 (0.92)	β^{KW} Interest rate	-0.09 (0.03)	-0.03 (0.08)

Notes: The table reports the estimated β^{CG} and β^{KW} in regression equations (33) and (34), respectively. For the model, we use model-simulated data, and report the median value across 10,000 simulations of 120 quarters. Standard errors are reported in parentheses.

current realization. Full-information rational expectations imply $\beta^{KW} = 0$, while the literature has documented that β^{KW} is negative and significant for a number of macro variables, including inflation, output growth and unemployment forecasts (Angeletos et al., 2020; Kohlhas and Walther, 2021). When current output or inflation are high, in aggregate, agents are too optimistic about the future path of output or inflation; they on average systematically "overshoot" the eventual truth.

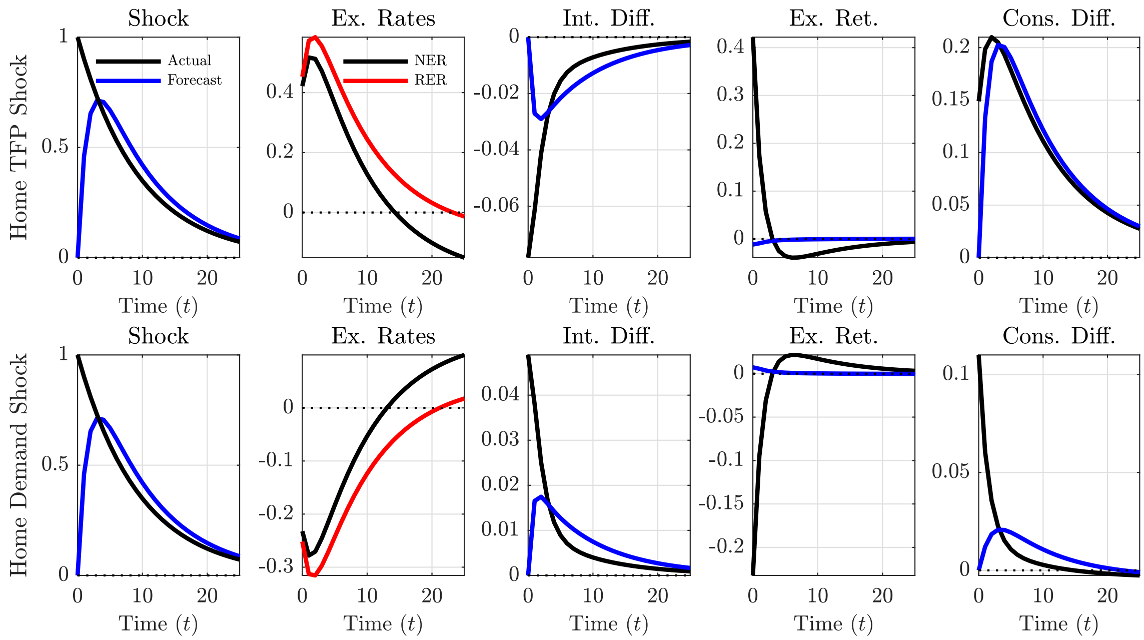
While the estimates of β^{CG} and β^{KW} may appear conflicting pieces of evidence, they can be rationalized by a model of expectations' formation that features for both under- and over-reaction, as recently shown by Angeletos et al. (2020). Therefore, we use the estimated values of β^{CG} and β^{KW} to discipline the extent of belief distortions of our GE model. Table 2 reports the estimates of β^{CG} and β^{KW} using data from *Survey of Professional Forecasters* from 1968:Q4 to 2019:Q4 for U.S. output and from 1981:Q2 to 2020:Q1 for U.S. inflation and interest rates. With the chosen κ_x and $\tilde{\rho}_x$, our model is able to reproduce the estimates of β^{CG} and β^{KW} for U.S. output and inflation (targeted) as well as for U.S. interest rates (not targeted).

4.2 Quantitative Results

We first describe the model dynamics using impulse responses. We then assess the model's ability to reproduce several moments of interest of exchange rates, macro variables, and the properties of their forecast errors.

Exchange rate dynamics in general equilibrium The first row of Figure 4 depicts the equilibrium time-t responses to a persistent home TFP increase, along with agents' forecasts from the previous period, $t-1$. Because agents hold biased beliefs

Figure 4: Impulse Responses to TFP and Demand Shocks



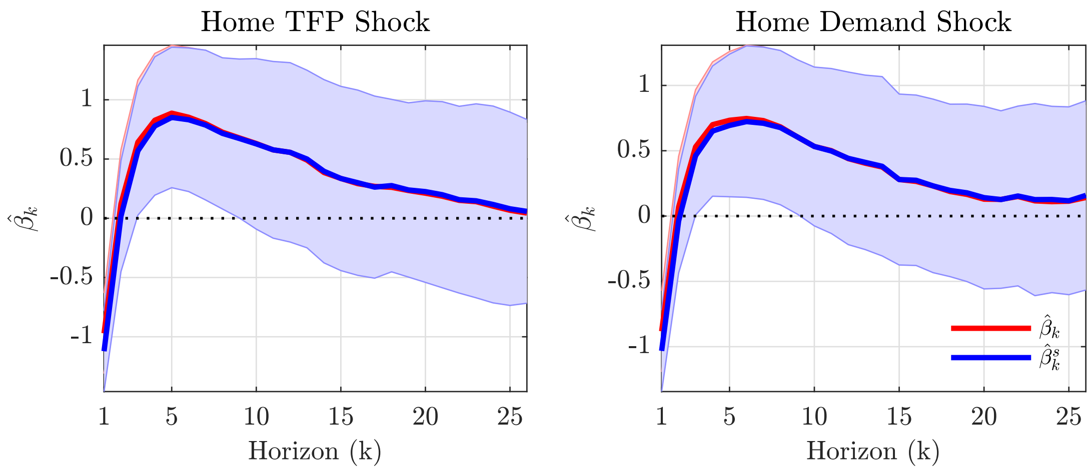
Notes: The first row reports the IRFs to a TFP shock, while the second row reports the IRFs to a demand shock. The x-axis denotes quarters from the shock, starting at 0.

about the TFP process, they initially under-predict the future path of TFP, while subsequently over-predict it. These belief misperceptions shape the equilibrium responses of all endogenous variables in the model.

Following the observed TFP increase, under incomplete markets the consumption differential increase driven by the increase in the wage differential. Because of nominal rigidities, the required downward adjustment of home prices is sluggish, resulting in an equilibrium decline in inflation and in the nominal interest rate differential. As a result, both nominal and real exchange rates depreciate on impact.

Initially, agents' forecasts systematically under-predict the subsequent, persistent, equilibrium decline in the interest rate differential, and thus they gradually revise downward their forecasts of its future path. Agents positive forecast errors on the interest rate differentials generate positive *ex-post* excess foreign-currency returns in the short run, accounting for the initial depreciation of the exchange rates. After around 5 quarters, agents' forecasts of future TFP become excessively optimistic leading to the opposite path of forecast errors and forecast revisions. The resulting over-reaction accounts for the sharp peak exchange rate depreciation and the observed reversal in ex-post excess returns. The joint responses of interest rate differentials and ex-post

Figure 5: UIP Regression Results (GE Model)



Notes: For both home TFP and home demand shocks, the figure reports the median estimates of β_k and β_k^s in equations (3) and (7), respectively, along with their 90% confidence interval, using 10,000 simulations of 120 quarters. Each horizon is a quarter.

excess returns following TFP shocks give rise to the empirically-relevant pattern of UIP deviations. Figure 5 reports the estimated β_k and β_k^s in equations (3) and (7), respectively, using model-simulated data. Figure 5 reveals that excess return predictability in the model is driven primarily by endogenous forecast errors rather than time-varying liquidity premia, in line with the empirical evidence in Section 1.

Figure 4 and Figure 5 also show that interest rates and excess returns dynamics following a demand shock also delivers the patterns of predictability of UIP deviations observed in the data. Thus, multiple fundamental shocks reproduce the empirically-consistent predictability of UIP deviations in the GE model with belief distortions.

Open-economy business-cycle moments We assess the ability of our GE model to reproduce several empirical moments of interest in open-economy macroeconomics. Panel A of Table 3 reports a number of exchange rate moments, both in the data and in the model. We estimate the moments for the US against the PPP-weighted sum of Canada, France, Germany, Italy, Japan and UK from 1978:Q3 to 2007:Q4. We report the moments for both the baseline model and for a version of the model without time-varying liquidity services (i.e., $\alpha = 0$).

The baseline model produces satisfactory results along many of the dimensions considered. As in the data, equilibrium nominal exchange rate changes display a low autocorrelation, and are significantly more volatile than macro variables. The model

Table 3: Exchange Rate Related Moments

Moments	Data	Model $\alpha = 0$	Model $\alpha = 0.15$	Moments	Data	Model $\alpha = 0$	Model $\alpha = 0.15$
A. Exchange Rates and Macro Variables				*B. Excess Returns Predictability*			
$\rho(\Delta s_t)$	0.16	0.29	0.28	β_1	-1.23 (0.90)	-0.81 (0.16)	-0.88 (0.20)
$\sigma(\Delta s_t)/\sigma(\Delta y_t)$	6.26	3.05	3.12	β_1^s	-1.74 (0.92)	-0.81 (0.16)	-1.04 (0.20)
$\rho(q_t)$	0.95	0.95	0.94	β_1^ξ	-0.46 (0.26)	0.00 (0.00)	-0.15 (0.00)
$\rho(\Delta q_t, \Delta s_t)$	0.99	1.00	1.00	Fama R^2	0.02	0.06	0.05
$\rho(\Delta q_t, \Delta(c_t - c_t^\star))$	-0.17	-0.19	-0.19	$\sum_k \beta_k^q$	> 0	> 0	> 0
$\sigma(r_t - r_t^\star)/\sigma(\Delta s_t)$	0.12	0.31	0.27				
$\rho(r_t - r_t^\star)$	0.77	0.72	0.71	*C. Stavrakeva-Tang Regressions*			
$\rho(\Delta c_t, \Delta c_t^\star)$	0.18	-0.05	-0.06	β^{ST}	0.17 (0.02)	0.47 (0.08)	0.48 (0.08)

Notes: For the model, each entry is the median value of moments across 10,000 simulations of 120 quarters. We report the moments for both the baseline model and for a specification without time-varying liquidity services ($\alpha = 0$). Standard errors are reported in parentheses (when relevant).

also reproduces the observed high autocorrelation of real exchange rates and their near-one correlation with the nominal exchange rate (PPP puzzle), as well as a negative correlation between the real exchange rate and consumption (Backus-Smith Puzzle). It is worth noting that while our mechanism relies on interest rate belief distortion, it does so within a model that features the empirically-relevant volatility and auto-correlation properties of equilibrium interest rates. A limitation of the model is that it does not deliver a positive international comovement in consumption, which could be addressed by allowing for correlated shocks across countries.

Expected Excess Currency Returns and Forecast Errors Our explanation of exchange rate puzzles does not rely on risk or liquidity premia. Indeed, Panel B of Table 3 shows that our model explains these puzzles even when $\alpha = 0$. Nevertheless, the presence of liquidity premia ($\alpha > 0$) allows us to examine the interaction between *expected* excess returns, forecast errors and their predictability.

Equation (6) highlights that the Fama β, i.e. β_1, can result from predictable forecast errors and/or predictable expected excess returns, that is $\beta_1 = \beta_1^s - \beta_1^\xi$. Section 1 documented that in advanced economies β_1^s drives the variation in β_1; for this reason, we primarily focused on the determinants of β_1^s. However, it is worth noting that one typically finds $\beta_1^\xi < 0$, although small and marginally significant. High interest rate differentials are associated with positive *ex-ante* excess returns on the foreign currency, even though *ex post* they turn out to be negative. Since the baseline model has both belief distortions and subjective deviations from UIP (due to time-varying liquidity yields) we can confront it with this fact.[25]

Panel B of Table 3 reports the estimates of β_1, β_1^s and β_1^ξ both in the data and in the baseline model. The model produces estimates of the overall Fama β and its components that align closely with the data, both qualitatively and quantitatively. When the home interest rate is high agents *expect* negative excess returns on the home currency ($\beta_1^\xi < 0$). In fact, when the interest rate is high liquid assets that can substitute for money become more valued for their liquidity services and so must pay a higher liquidity return. To the contrary, because agents hold distorted beliefs on the underlying fundamental processes, they initially under-estimate the future path of interest rate differentials and liquidity premia, which accounts for short-run forecast errors ($\beta_1^s < 0$) and overall high *realized* excess returns on the home currency ($\beta_1 < 0$).[26]

Exchange rate beliefs We now test whether the exchange rate expectations implied by our microfoundation are consistent with previous evidence on exchange rate beliefs. Using *Consensus Economics* exchange rate forecast data, Stavrakeva and Tang (2020b) show that one can predict exchange rate forecast errors using past exchange rate depreciations. Their results suggest that past depreciations are associated with the forecasters' belief that the currency will appreciate by more than if forecasters' expectations were rational.

Consider a regression of exchange rate forecast errors on past depreciations:

$$s_{t+1} - \mathrm{E}_t^s \, s_{t+1} = \alpha^{ST} + \beta^{ST}(s_t - s_{t-1}) + \varepsilon_{t+1}^{ST}. \tag{35}$$

Rational expectations imply $\beta^{ST} = 0$, while Stavrakeva and Tang (2020b) find that β^{ST} is positive and significant for almost all currency pairs.[27] This implies that forecasters

[25] For a broader discussion of survey-based risk premia, see Kalemli-Ozcan and Varela (2021).

[26] In this model with liquidity premia agents make forecast errors on both the interest rate differential under UIP and the liquidity premia. In fact, the predictability of forecast errors results from systematic revisions of both the UIP-consistent exchange rate and its liquidity premium component.

[27] Stavrakeva and Tang (2020b) document these patterns for forecast horizons of 1, 3, 12, and 24 months. In Stavrakeva and Tang's (2020b) regressions the relevant estimated coefficient is negative as their definition of forecast error is opposite relative to our notation.

underweight past exchange rate changes relative to rational expectations. We estimate regression (35) at quarterly frequency using *Consensus Economics* data from 1990:M1 to 2019:M7 for the currency pairs considered in Section 1. Our panel estimate of β^{ST}, which are in line with Stavrakeva and Tang (2020b), are reported in Panel C of Table 3, along with the moments produced by model-simulated data.

The model very naturally reproduces this dimension of exchange rate forecast error predictability. In fact, $\mathrm{Cov}(s_t - s_{t-1}, s_{t+1} - \mathrm{E}_t^s\, s_{t+1}) > 0$ during both phases of expectations' under and over-reaction. When agents under-predict the future path of interest differentials, the exchange rate tends to appreciate *while* they under-predict the ensuing appreciation. When agents over-predict the future path of interest differentials, the exchange rate tends to depreciate *while* agents under-predict the resulting rates of depreciation. Our set of belief distortions is thus able to reproduce several forecast error predictability properties of both macro variables (Table 2) and exchange rates.

4.3 Monetary Policy and the New Fama Puzzle

In this paper, UIP violations arise due to agents' biased expectations of underlying fundamentals. In GE, biased beliefs about fundamentals result in forecast errors about endogenous variables including interest rates, output, and inflation. As a result, the patterns of UIP violations critically depend on agents' biases *and* the structural relationships implied by the model. In particular, while fundamentals follow exogenous processes, the dynamic properties of interest rates and ensuing forecast errors are endogenous to monetary policy in our GE framework.

Recent evidence re-examining the UIP puzzle allows to exploit this particular insight to test the main mechanism of our model. Specifically, Bussiere et al. (2018) document that the Fama β turns positive after the Great Recession, in conjunction with a change in the conduct of monetary policy, for 8 advanced countries' exchange rates against the US dollar. That is, in this period a rise in the domestic interest rates relative to the Federal Funds rate predicts a depreciation of the country's currency against the US dollar in excess of UIP.[28] We find a similar pattern of β_1 in our sample of advanced economies, as reported in the first column of Table 4. Crucially, Bussiere et al. (2018) document that the "new Fama puzzle" – the switch in the sign of the β_1 relative to the pre-crisis sample – is mostly due to the change in the sign of the covariance between

[28] Along these lines, Engel et al. (2021) document that interest rate differentials fail to robustly predict excess dollar returns in samples that include post-2008 data.

forecast errors and interest rate differentials. The endogeneity of this covariance to policy, captured by the coefficient β_1^s, is a distinctive feature of our GE framework. Thus, verifying whether the model generates different β_1^s's in different policy regimes constitutes a natural test of our mechanism.

To operationalize this test, we depart from the baseline policy regime, that is an equilibrium that features active monetary and passive fiscal policies (AM/PF), where the monetary authority reacts strongly to inflation, while the fiscal authority adjusts taxes to fully fund its debt.[29] Following Bianchi and Melosi (2017), we interpret the post-2008 period as a regime of passive monetary and active fiscal policies (PM/AF).[30] In this regime, the fiscal authority does not adjust taxes sufficiently to stabilize debt, and deficits are financed by a "passive" monetary authority ($\phi_\pi < 1$) allowing inflation to rise and inflate debt away (Leeper, 1991; Bianchi and Melosi, 2014).

The government budget constraint is reported in equation (28). As in Bianchi and Melosi (2017), our general formulation of the tax rule allows the tax-to-output ratio, $\tau_t = \frac{P_t T_t}{P_{Ht} Y_{Ht}}$, to respond to the debt-to-output ratio, $\mathcal{B}_{H,t-1} = \frac{P_{t-1} B_{H,t}}{P_{H,t-1} Y_{H,t-1}}$, and to the the output gap, $\hat{y}_t - \hat{y}_t^n$. The rule also allows for a certain degree of tax smoothing:

$$\tilde{\tau}_t = \rho_\tau \tilde{\tau}_t + (1 - \rho_\tau) \left[\delta_b \tilde{\mathcal{B}}_{H,t-1} + \delta_y (\hat{y}_t - \hat{y}_t^n) \right], \tag{36}$$

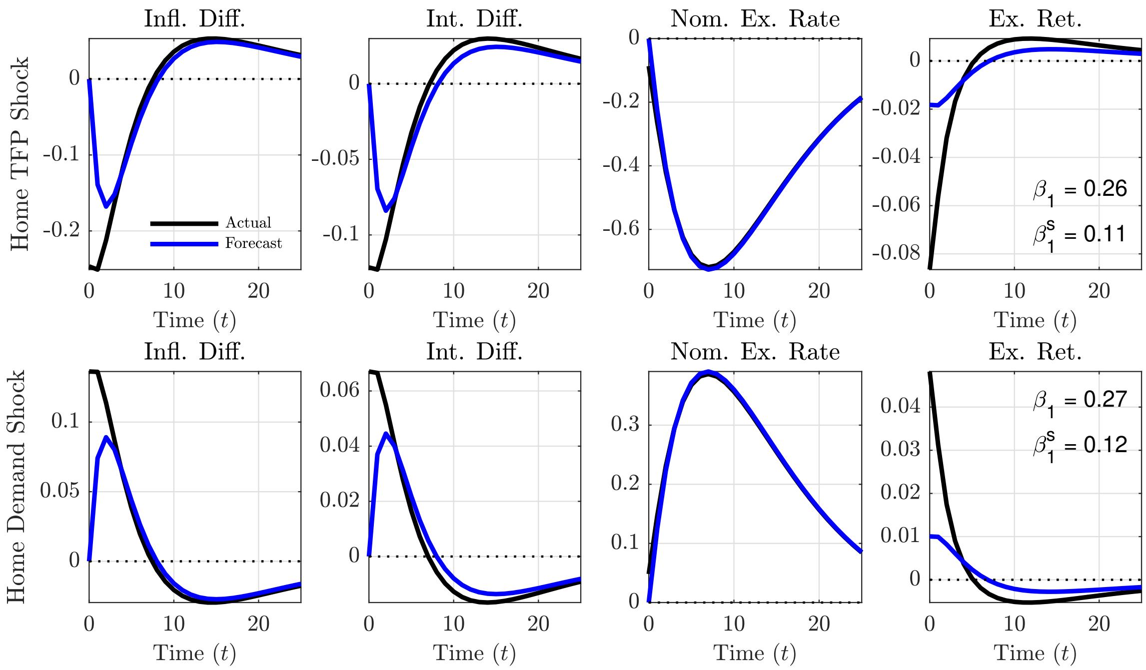
where a tilde over a variable denotes the level of the variable in deviations from the steady state. In our baseline PM/AF regime, we assume that tax revenues adjust to changes in output in order to keep the tax-to-output ratio unchanged. We thus set $\phi_\pi = 0.5$, $\rho_\tau = \delta_b = \delta_y = 0$, while the remaining parameter are calibrated as in Table 1. Below, we explore different calibrations of the parameters in the tax rule.

Figure 6 depicts the impulse responses to home TFP and home demand shocks in the PM/AF regime. The figure reveals that interest rates and exchange rates comove differently than in the AM/PF regime. In fact, under PM/AF high interest rate differentials are associated with positive rates of depreciations in excess of UIP. That is, high interest rates predict positive foreign excess currency returns, and importantly, more than what is required as a liquidity premium. As a result, $\beta_1 > 0$ and $\beta_1^s > 0$. This

[29] While this configuration reasonably characterizes the experience of advanced economies pre-2008, central banks could not arguably react to inflation sufficiently strongly post-2008 when the interest rate was constrained by the effective lower bound.

[30] In their structural estimation exercise, Bianchi and Melosi (2017) interpret the post-2008 period as an extreme version of a PM/AF regime to distinguish this period from the pre-Volcker period of PM/AF that occurred in their sample. This distinction is not needed for our purposes.

Figure 6: Impulse Responses Under Passive Monetary and Active Fiscal Policies



Notes: The figure reports the IRFs to a positive home TFP shock (first row) and a positive home demand shock (second row) under the PM/AF regime. For both TFP and demand shocks, the figure reports the median estimates of β_1 and β_1^s in equations (3) and (7), respectively, along with the standard deviation in parentheses, using 10,000 simulations of 120 quarters.

contrast sharply with the AM/PF regime, where positive interest rates differentials are associated with subsequent negative foreign excess currency returns (Figure 4).

The intuition for the switch in β_1 and β_1^s is simple, and follows from the properties of interest rate forecast revisions in the PM/AF regime in the presence of under-reaction. In the PM/AF regime negative inflation systematically predicts future positive inflation. In fact, negative inflation increases the real value of government debt which, absent a strong fiscal adjustment, will be stabilized by future positive rates of inflation. Under a Taylor rule (albeit with $\phi_\pi < 1$), nominal interest rates inherit the dynamics properties of inflation: high interest rates today predict *low* interest rates in the future. Moreover, whenever the interest rate today is higher than agents' prior, agents systematically revise their forecast of the cumulative path of the interest rate *downwards*.[31] Because agents under-react to persistent changes in fundamentals, a shock that increases interest rates today will lead agents to be surprised for several

[31] Appendix E.4 proves this analytically when tax revenues are constant, while Figure 6 confirms that this applies also to the more realistic case in which tax revenues adjust to changes in output.

Table 4: UIP Regressions Across Regimes

	Data	Model			
Pre-2008 β_1	-1.26 (0.90)	AM/PF	-0.88 (0.20)	-0.88 (0.20)	-0.88 (0.20)
Post-2008 β_1	1.93 (2.30)	PM/AF	0.26 (0.02)	0.55 (0.17)	0.56 (0.17)
Tax rule (eq. (36))			$\rho_\tau = 0$ $\delta_y = 0$	$\rho_\tau = 0$ $\delta_y = 0.25$	$\rho_\tau = 0.45$ $\delta_y = 0.25$

Notes: The table reports the estimates of β_1 in regression equation (2), in both data and model. For the model, the pre-2008 values are computed using the baseline AM/PF regime described above, while the post-2008 values are computed using the PM/AF regime under different calibrations of the parameters of the fiscal block. Under AM/PF $\phi_\pi = 2.50$, while under PM/AF $\phi_\pi = 0.50$. The remaining parameter are calibrated as in Table 1.

periods by high interest rates. These positive surprises induce agents to *continuously* revise the forecast of the interest rate path *downwards* resulting in an *over-depreciation* of the exchange rate in the short run.

This mechanism generates positive β_1 and β_1^s and does not make reference to a particular type of shock, thus applying equally to supply or demand disturbances.

Table 4 shows that the Fama coefficient β_1 from our model simulations rationalizes the "old" and "new" Fama puzzle in our baseline calibration. The table confirms that the result is robust to alternative parameterizations of the tax rule taken from the estimates of Bianchi and Melosi (2017). Indeed, a response of taxes to the output gap either with or without smoothing affects the positive β_1 in the PM/AF regime only quantitatively but leaves the sign unchanged.

From this exercise we conclude that excess currency returns in our GE model are endogenous to the policy regime in a way that jointly rationalizes the evidence about the direction of UIP deviations before *and* after the financial crisis. As the empirical switch has been attributed to changes in the correlation between forecast errors and interest rates, we interpret the ability of our model to reproduce these facts as directly supportive of the key mechanism that we proposed. In our model, both β_1 and β_1^s are naturally positive when monetary policy is passive, due to the joint properties of inflation, interest rates and forecast error dynamics that obtain in the PM/AF regime.[32]

[32] In a model with noisy signals and over-extrapolation about the interest rate differential, Valente et al. (2021) offer a distinct explanation of the New Fama Puzzle. Their explanation is based on a reduction in the dispersion of beliefs about interest rates. In their model, lower beliefs dispersion is associated with less under-reaction and can lead to a positive Fama β if over-reaction dominates.

5 Conclusions

The notion that exchange rate fluctuations are partly driven by incorrect expectations about macroeconomic fundamentals is prevalent among academics and practitioners. Yet, *if* and *how* expectational errors influence exchange rate dynamics is not well understood. Following a long tradition in macroeconomics and international finance, we used survey expectations to document new properties of exchange rate forecast errors. Our evidence revealed that expectations under-reaction and over-reaction coexist and operate at different horizons, and this property is responsible for the predictability of excess currency returns recently documented by Engel (2016) and Valchev (2020).

The coexistence of under and over-reaction of exchange rate beliefs implies that simple behavioral theories of exchange rate determination inevitably fall short of explaining the whole dynamics of forecast errors. We proposed a theory of shock misperceptions and over-extrapolation that rationalizes the observed dynamics of exchange rate expectations. Under-reaction arises because agents overstate the importance of transitory shocks. Over-reaction stems from the fact that investors perceive persistent shocks to be more autocorrelated than they actually are. In this context, under-reaction naturally dominates in the short-run while over-reaction prevails in the longer-run. We showed that these parsimonious ingredients not only explains several exchange rate puzzles but also, once embedded in a standard open-economy general-equilibrium model, make clear predictions about forecast errors of interest rates and other key macroeconomic indicators that are supported by the data.

Our findings suggest that the predictability of macro and exchange rate forecast errors are driven by the same behavioral biases. Studying the interaction between these biases and policy is crucial to understand exchange rate dynamics. In fact, we have shown that this particular interaction provides a simple, empirically-supported, explanation of the new Fama puzzle that emerged after the financial crises. Future research exploring these interactions in other financial markets seems promising.

We provided new disciplining evidence for models of exchange rate expectations and identified the set of belief distortions that are necessary to make sense of the data. Like much of the finance literature, we remained silent on the specific behavioral biases that operate behind the belief distortions included in our model. A natural direction for future work involves identifying the origins of such behavioral biases. Understanding the nature of these belief distortions is all the more important since contemporaneous work by Angeletos et al. (2020) and Bianchi et al. (2020) suggest that these non-monotonic patterns of forecast errors are pervasive macroeconomic phenomena.

References

Alvarez, F., A. Atkeson, and P. J. Kehoe (2009). Time-varying risk, interest rates, and exchange rates in general equilibrium. *The Review of Economic Studies 76*(3), 851–878.

Angeletos, G.-M., Z. Huo, and K. A. Sastry (2020). Imperfect Macroeconomic Expectations: Evidence and Theory. In *NBER Macroeconomics Annual 2020, volume 35*, NBER Chapters. National Bureau of Economic Research, Inc.

Bacchetta, P., E. Mertens, and E. Van Wincoop (2009). Predictability in financial markets: What do survey expectations tell us? *Journal of International Money and Finance 28*(3), 406–426.

Bacchetta, P. and E. Van Wincoop (2006). Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle? *American Economic Review 96*(3), 552–576.

Bacchetta, P. and E. Van Wincoop (2010). Infrequent Portfolio Decisions: A Solution to the Forward Discount Puzzle. *American Economic Review 100*(3), 870–904.

Bacchetta, P. and E. van Wincoop (2021). Puzzling exchange rate dynamics and delayed portfolio adjustment. *Journal of International Economics 131*, 103460.

Backus, D. K., P. J. Kehoe, and F. E. Kydland (1994). Dynamics of the Trade Balance and the Terms of Trade: The J-Curve? *The American Economic Review 84*(1), pp. 84–103.

Bansal, R. and I. Shaliastovich (2012). A long-run risks explanation of predictability puzzles in bond and currency markets. *Review of Financial Studies*, hhs108.

Barberis, N., A. Shleifer, and R. Vishny (1998). A model of investor sentiment. *Journal of Financial Economics 49*(3), 307–343.

Benigno, P. (2009). Price Stability with Imperfect Financial Integration. *Journal of Money, Credit and Banking 41*(s1), 121–149.

Bianchi, F., S. C. Ludvigson, and S. Ma (2020). Belief distortions and macroeconomic fluctuations. Working Paper 27406, National Bureau of Economic Research.

Bianchi, F. and L. Melosi (2014). Dormant shocks and fiscal virtue. *NBER Macroeconomics Annual 28*(1), 1–46.

Bianchi, F. and L. Melosi (2017). Escaping the Great Recession. *American Economic Review 107*(4), 1030–58.

Bilson, J. F. O. (1981). The "Speculative Efficiency" Hypothesis. *The Journal of Business 54*(3), 435–451.

Blanchard, O. J., J.-P. L'Huillier, and G. Lorenzoni (2013). News, noise, and fluctuations: An empirical exploration. *American Economic Review 103*(7), 3045–70.

Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer (2020). Overreaction in Macroeconomic Expectations. *American Economic Review 110*(9), 2748–2782.

Bordalo, P., N. Gennaioli, R. L. Porta, and A. Shleifer (2019). Diagnostic Expectations and Stock Returns. *Journal of Finance 74*(6), 2839–2874.

Bordalo, P., N. Gennaioli, and A. Shleifer (2018). Diagnostic Expectations and Credit Cycles. *Journal of Finance 73*(1), 199–227.

Bunsupha, S. (2018). Extrapolative Beliefs and Exchange Rate Markets.

Burnside, C., B. Han, D. Hirshleifer, and T. Y. Wang (2011). Investor Overconfidence and the Forward Premium Puzzle. *Review of Economic Studies 78*(2), 523–558.

Bussiere, M., M. D. Chinn, L. Ferrara, and J. Heipertz (2018). The New Fama Puzzle. Working Paper 24342, National Bureau of Economic Research.

Candian, G. (2019). Information frictions and real exchange rate dynamics. *Journal of International Economics 116*, 189–205.

Chari, V. V., P. J. Kehoe, and E. R. McGrattan (2002). Can Sticky Price Models Generate Volatile and Persistent Real Exchange Rates? *The Review of Economic Studies 69*(3), pp. 533–563.

Chernov, M. and D. Creal (2020). The PPP View of Multihorizon Currency Risk Premiums. *The Review of Financial Studies 34*(6), 2728–2772.

Chinn, M. and J. Frankel (2019). A third of a century of currency expectations data: the carry trade and the risk premium.

Coibion, O. and Y. Gorodnichenko (2015). Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts. *American Economic Review 105*(8), 2644–2678.

Colacito, R. and M. M. Croce (2013). International asset pricing with recursive preferences. *The Journal of Finance 68*(6), 2651–2686.

Dahlquist, M. and J. Pénasse (2021). The missing risk premium in exchange rates. *Journal of Financial Economics*.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam (1998). Investor Psychology and Security Market Under- and Overreactions. *Journal of Finance 53*(6), 1839–1885.

Dornbusch, R. (1976). Expectations and Exchange Rate Dynamics. *Journal of Political Economy 84*(6), 1161–1176.

Driscoll, J. C. and A. C. Kraay (1998). Consistent Covariance Matrix Estimation With Spatially Dependent Panel Data. *The Review of Economics and Statistics 80*(4), 549–560.

Eichenbaum, M. and C. L. Evans (1995). Some Empirical Evidence on the Effects of Shocks to Monetary Policy on Exchange Rates. *The Quarterly Journal of Economics 110*(4), 975–1009.

Engel, C. (2014). Exchange Rates and Interest Parity. In G. Gopinath, . Helpman, and K. Rogoff (Eds.), *Handbook of International Economics*, Volume 4 of *Handbook of International Economics*, Chapter 8, pp. 453–522. Elsevier.

Engel, C. (2016). Exchange Rates, Interest Rates, and the Risk Premium. *American Economic Review 106*(2), 436–474.

Engel, C., E. Kazakova, M. Wang, and N. Xiang (2021). A Reconsideration of the Failure of Uncovered Interest Parity for the U.S. Dollar. Working Paper 28420, National Bureau of Economic Research.

Engel, C., N. C. Mark, and K. D. West (2008). Exchange Rate Models Are Not As Bad As You Think. In *NBER Macroeconomics Annual 2007, Volume 22*, NBER Chapters, pp. 381–441. National Bureau of Economic Research, Inc.

Engel, C. and S. P. Y. Wu (2020). Liquidity and Exchange Rates: An Empirical Investigation.

Fama, E. F. (1984). Forward and spot exchange rates. *Journal of Monetary Economics 14*(3), 319 – 338.

Farhi, E. and X. Gabaix (2015). Rare disasters and exchange rates. *The Quarterly Journal of Economics*.

Frankel, J. A. and K. A. Froot (1987). Using Survey Data to Test Standard Propositions Regarding Exchange Rate Expectations. *American Economic Review 77*(1), 133–153.

Froot, K. A. and J. A. Frankel (1989). Forward discount bias: Is it an exchange risk premium? *The Quarterly Journal of Economics 104*(1), 139–161.

Gabaix, X. (2020). A Behavioral New Keynesian Model. *American Economic Review 110*(8), 2271–2327.

Gabaix, X. and M. Maggiori (2015). International liquidity and exchange rate dynamics. *The Quarterly Journal of Economics 130*(3), 1369–1420.

Gennaioli, N., Y. Ma, and A. Shleifer (2016). Expectations and investment. *NBER Macroeconomics Annual 30*(1), 379–431.

Gourinchas, P.-O. and A. Tornell (2004). Exchange rate puzzles and distorted beliefs. *Journal of International Economics 64*(2), 303–333.

Granziera, E. and M. Sihvonen (2020). Bonds, currencies and expectational errors. Working Paper 2020/3, Norges Bank.

Greenwood, R. and A. Shleifer (2014). Expectations of returns and expected returns. *The Review of*

Financial Studies 27(3), 714–746.

Guo, H. and J. A. Wachter (2019). "Superstitious" Investors.

Hamilton, J. (1994). *Time Series Analysis.* Princeton University Press.

Hetting, T., G. Muller, and M. Wolf (2021). Delayed Overshooting: The Case for Informational Rigidities.

Hong, H. and J. C. Stein (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of Finance 54*(6), 2143–2184.

Ilut, C. (2012). Ambiguity Aversion: Implications for the Uncovered Interest Rate Parity Puzzle. *American Economic Journal: Macroeconomics 4*(3), 33–65.

Itskhoki, O. and D. Mukhin (2021). Exchange Rate Disconnect in General Equilibrium. *Journal of Political Economy*.

Jeanne, O. and A. K. Rose (2002). Noise Trading and Exchange Rate Regimes. *The Quarterly Journal of Economics 117*(2), 537–569.

Jiang, Z., A. Krishnamurthy, and H. Lustig (2021). Foreign Safe Asset Demand and the Dollar Exchange Rate. *The Journal of Finance 76*(3), 1049–1089.

Jordà, O. (2005). Estimation and Inference of Impulse Responses by Local Projections. *American Economic Review 95*(1), 161–182.

Kalemli-Ozcan, S. and L. Varela (2021). Five Facts about the UIP Premium.

Kohlhas, A. N. and A. Walther (2021). Asymmetric attention. *American Economic Review 111*(9), 2879–2925.

Leeper, E. M. (1991). Equilibria under 'active' and 'passive' monetary and fiscal policies. *Journal of Monetary Economics 27*(1), 129–147.

Lustig, H., A. Stathopoulos, and A. Verdelhan (2019, December). The term structure of currency carry trade risk premia. *American Economic Review 109*(12), 4142–77.

Molavi, P., A. Tahbaz-Salehi, and A. Vedolin (2021). Model Complexity, Expectations, and Asset Prices. Technical report, National Bureau of Economic Research.

Nagel, S. (2016). The Liquidity Premium of Near-Money Assets. *The Quarterly Journal of Economics 131*(4), 1927–1971.

Obstfeld, M. and K. Rogoff (1995). Exchange Rate Dynamics Redux. *Journal of Political Economy 103*(3), pp. 624–660.

Stavrakeva, V. and J. Tang (2020a). A Fundamental Connection: Survey-based Exchange Rate Decomposition.

Stavrakeva, V. and J. Tang (2020b). Deviations from FIRE and Exchange Rates: a GE Theory of Supply and Demand.

Turnovsky, S. J. (1985). Domestic and foreign disturbances in an optimizing model of exchange-rate determination. *Journal of International Money and Finance 4*(1), 151–171.

Uhlig, H. F. et al. (1995). A toolkit for analyzing nonlinear dynamic stochastic models easily.

Valchev, R. (2020). Bond Convenience Yields and Exchange Rate Dynamics. *American Economic Journal: Macroeconomics*.

Valente, J. P., K. Vasudevan, and T. Wu (2021). The role of beliefs in asset prices: Evidence from exchange rates.

Verdelhan, A. (2010). A habit-based explanation of the exchange rate risk premium. *The Journal of Finance 65*(1), 123–146.

Appendix

A Data

- Forward and spot exchange rates; *Reuters/WMR and Barclays, retrieved from Datastream*
  - *Frequency*: Monthly
  - *Currencies pairs*: CAD/USD, DEK/USD, DEM-EUR/USD, JPY/USD, NOK/USD, SEK/USD, CHF/USD, GBP/USD.
  - *Interest rate differentials:* We construct interest rate differentials from the Covered Interest Parity (CIP):[33]

$$\frac{F_t}{S_t} = \frac{1 + r_t}{1 + r_t^*}$$

- Exchange rates forecasts; *Consensus Economics Foreign Exchange Consensus Forecasts*
  - *Frequency*: Monthly
  - *Horizon*: 3-month ahead
  - *Currencies pairs*: CAD/USD, DEK/USD, DEM-EUR/USD, JPY/USD, NOK/USD, SEK/USD, CHF/USD, GBP/USD.
- Interest rates and interest rates forecasts; *FX4Casts*
  - Three-month LIBOR rates and three-month LIBOR rates forecast
  - *Frequency:* Monthly
  - *Countries:* Canada, Germany-EZ, Switzerland, UK, Japan
- U.S. output, U.S. inflation, U.S. interest rates forecasts; *Survey of Professional Forecasters*
  - *Frequency:* Quarterly

[33] This is the standard practice in the literature because the data on forward contracts are more widely available than data on short-term interest rates. When both series are available they are virtually identical.

- Empirical moments in Table 3; *Various sources, retrieved from FRED*
  - *Frequency:* Quarterly
  - *Countries:* Home: United States; Foreign: PPP-weighted sum of Canada, France, Germany, Italy, Japan, UK.

B Proofs

B.1 Proof of Proposition 1

From the definitions of the perceived interest rate process (10)-(11) it follows that $\mathrm{E}_t^s x_{t+j} = \mathrm{E}_t^s z_{t+j} = \tilde{\rho}^{j-1} \mathrm{E}_t^s z_{t+1} = \tilde{\rho}^{j-1} \mathrm{E}_t^s x_{t+1}$. Substituting this result in equation (14) and noting that $\mathrm{E}_t^s x_t = x_t$ we obtain equation (15).

B.2 Proof of Proposition 2

The coefficients β_k's are defined as:

$$\beta_{k+1} = \frac{\mathrm{Cov}(s_{t+k+1} - \mathrm{E}_{t+k}^s s_{t+k+1}, x_t)}{\mathrm{Var}(x_t)}. \tag{A.1}$$

From the solution of the exchange rate in (15) we obtain the following expression for the one-step ahead forecast error:

$$\begin{aligned} s_{t+1} - \mathrm{E}_t^s s_{t+1} &= -x_{t+1} - \frac{1}{1-\tilde{\rho}} \mathrm{E}_{t+1}^s z_{t+2} + (\mathrm{E}_t^s x_{t+1} + \frac{1}{1-\tilde{\rho}} \mathrm{E}_t^s \mathrm{E}_{t+1}^s z_{t+1}), \\ &= -x_{t+1} - \frac{\tilde{\rho}}{1-\tilde{\rho}} \mathrm{E}_{t+1}^s z_{t+1} + (\mathrm{E}_t^s z_{t+1} + \frac{\tilde{\rho}}{1-\tilde{\rho}} \mathrm{E}_t^s z_{t+1}), \\ &= -x_{t+1} - \frac{\rho_s}{1-\tilde{\rho}} (\mathrm{E}_{t+1}^s z_{t+1} - \mathrm{E}_t^s z_t). \end{aligned} \tag{A.2}$$

The Kalman filter implies that the beliefs about z_t evolve according to:

$$\mathrm{E}_t^s z_t = (1-\kappa)\tilde{\rho}\, \mathrm{E}_{t-1}^s z_{t-1} + \kappa x_t. \tag{A.3}$$

We can use this expression to rewrite the forecast error in (A.2) as:

$$s_{t+1} - \mathrm{E}_t^s s_{t+1} = \frac{1-(1-\kappa)\tilde{\rho}}{1-\tilde{\rho}} (-x_{t+1} + \tilde{\rho}\, \mathrm{E}_t^s z_t) \tag{A.4}$$

Also, solving equation (A.3) backward yields:

$$\mathrm{E}_t^s\, z_t = \kappa \sum_{i=0}^{\infty} \left[(1-\kappa)\tilde{\rho}\right]^i x_{t-i} \tag{A.5}$$

Substituting this expression in (A.4) we finally obtain:

$$s_{t+1} - \mathrm{E}_t^s\, s_{t+1} = \frac{1-(1-\kappa)\tilde{\rho}}{1-\tilde{\rho}} \left(-x_{t+1} + \tilde{\rho}\kappa \sum_{i=0}^{\infty} \left[(1-\kappa)\tilde{\rho}\right]^i x_{t-i} \right) \tag{A.6}$$

We now proceed to computing the β_{k+1}'s. Let us begin by β_1:

$$\begin{aligned}
\beta_1 &= \frac{\mathrm{Cov}(s_{t+1} - \mathrm{E}_t^s\, s_{t+1}, x_t)}{\mathrm{Var}(x_t)} \\
&= \frac{\frac{1-(1-\kappa)\tilde{\rho}}{1-\tilde{\rho}} \left(-\mathrm{Cov}(x_{t+1}, x_t) + \tilde{\rho}\kappa \sum_{i=0}^{\infty} \left[(1-\kappa)\tilde{\rho}\right]^i \mathrm{Cov}(x_{t-i}, x_t) \right)}{\mathrm{Var}(x_t)} \\
&= \frac{\frac{1-(1-\kappa)\tilde{\rho}}{1-\tilde{\rho}} \left(-\rho\, \mathrm{Var}(x_t) + \tilde{\rho}\kappa \sum_{i=0}^{\infty} \left[(1-\kappa)\tilde{\rho}\right]^i \rho^i\, \mathrm{Var}(x_t) \right)}{\mathrm{Var}(x_t)} \\
&= \frac{\frac{1-(1-\kappa)\tilde{\rho}}{1-\tilde{\rho}} \left(-\rho + \frac{\tilde{\rho}\kappa}{1-(1-\kappa)\tilde{\rho}\rho} \right) \mathrm{Var}(x_t)}{\mathrm{Var}(x_t)} = \frac{1-(1-\kappa)\tilde{\rho}}{1-\tilde{\rho}} \left(-\rho + \frac{\tilde{\rho}\kappa}{1-(1-\kappa)\tilde{\rho}\rho} \right).
\end{aligned}$$

Now, exploiting the stationarity of the model, we can calculate $\beta_2 = \frac{\mathrm{Cov}(s_{t+2} - \mathrm{E}_{t+1}^s\, s_{t+2}, x_t)}{\mathrm{Var}(x_t)}$ as $\frac{\mathrm{Cov}(s_{t+1} - \mathrm{E}_t^s\, s_{t+1}, x_{t-1})}{\mathrm{Var}(x_t)}$. Defining the composite parameter $a = \frac{1-(1-\kappa)\tilde{\rho}}{1-\tilde{\rho}}$,the numerator of this expression is:

$$\begin{aligned}
\mathrm{Cov}(s_{t+1} - \mathrm{E}_t^s\, s_{t+1}, x_{t-1}) &= a \left(-\mathrm{Cov}(x_{t+1}, x_{t-1}) + \tilde{\rho}\kappa \sum_{i=0}^{\infty} \left[(1-\kappa)\tilde{\rho}\right]^i \mathrm{Cov}(x_{t-i}, x_{t-1}) \right) \\
&= a\Big(-\rho^2\, \mathrm{Var}(x_t) + \tilde{\rho}\kappa(1-\kappa)\tilde{\rho} \sum_{i=0}^{\infty} \left[(1-\kappa)\tilde{\rho}\right]^i \mathrm{Cov}(x_{t-i-1}, x_{t-1}) + \\
&+ \tilde{\rho}\kappa\, \mathrm{Cov}(x_t, x_{t-1}) \Big) \\
&= a\Big(-\rho^2\, \mathrm{Var}(x_t) + \tilde{\rho}\kappa(1-\kappa)\tilde{\rho} \sum_{i=0}^{\infty} \left[(1-\kappa)\tilde{\rho}\right]^i \rho^i\, \mathrm{Var}(x_t) + \rho\tilde{\rho}\kappa\, \mathrm{Var}(x_t) \Big) \\
&= a\Big(-\rho^2 + \frac{\tilde{\rho}\kappa(1-\kappa)\tilde{\rho}}{1-(1-\kappa)\tilde{\rho}\rho} + \rho\tilde{\rho}\kappa \Big)\, \mathrm{Var}(x_t).
\end{aligned}$$

It follows that:

$$\beta_2 = \frac{1-(1-\kappa)\tilde{\rho}}{1-\tilde{\rho}}\Big(-\rho^2 + \frac{\tilde{\rho}\kappa(1-\kappa)\tilde{\rho}}{1-(1-\kappa)\tilde{\rho}\rho} + \rho\tilde{\rho}\kappa\Big).$$

Proceeding in a similar manner for β_3 one obtains the expression:

$$\beta_3 = \frac{1-(1-\kappa)\tilde{\rho}}{1-\tilde{\rho}}\Big(-\rho^3 + \frac{\tilde{\rho}\kappa(1-\kappa)^2\tilde{\rho}^2}{1-(1-\kappa)\tilde{\rho}\rho} + \tilde{\rho}\kappa(\rho^2 + (1-\kappa)\tilde{\rho}\rho)\Big).$$

$$\begin{aligned}
&\operatorname{Cov}(s_{t+1} - \operatorname{E}_t^s s_{t+1}, x_{t-2}) \\
&= a\left(-\operatorname{Cov}(x_{t+1}, x_{t-2}) + \tilde{\rho}\kappa \sum_{i=0}^{\infty} [(1-\kappa)\tilde{\rho}]^i \operatorname{Cov}(x_{t-i}, x_{t-2})\right) \\
&= a\Big(-\rho^3 \operatorname{Var}(x_t) + \tilde{\rho}\kappa(1-\kappa)^2\tilde{\rho}^2 \sum_{i=0}^{\infty} [(1-\kappa)\tilde{\rho}]^i \operatorname{Cov}(x_{t-i-2}, x_{t-2}) + \\
&+ \tilde{\rho}\kappa \operatorname{Cov}(x_t, x_{t-2}) + \tilde{\rho}\kappa(1-\kappa)\tilde{\rho}\operatorname{Cov}(x_{t-1}, x_{t-2})\Big) \\
&= a\Big(-\rho^3 \operatorname{Var}(x_t) + \tilde{\rho}\kappa(1-\kappa)^2\tilde{\rho}^2 \sum_{i=0}^{\infty} [(1-\kappa)\tilde{\rho}]^i \rho^i \operatorname{Var}(x_t) + \tilde{\rho}\kappa\rho^2 \operatorname{Var}(x_t) + \tilde{\rho}\kappa(1-\kappa)\tilde{\rho}\rho \operatorname{Var}(x_t)\Big) \\
&= a\Big(-\rho^3 + \frac{\tilde{\rho}\kappa(1-\kappa)^2\tilde{\rho}^2}{1-(1-\kappa)\tilde{\rho}\rho} + \tilde{\rho}\kappa(\rho^2 + (1-\kappa)\tilde{\rho}\rho)\Big) \operatorname{Var}(x_t).
\end{aligned}$$

It follows that the general pattern of the β_{k+1} can be written as:

$$\begin{aligned}
&a\Big(-\rho^{k+1} + \frac{\tilde{\rho}\kappa(1-\kappa)^k\tilde{\rho}^k}{1-(1-\kappa)\tilde{\rho}\rho} + \tilde{\rho}\kappa(\rho^k + (1-\kappa)\tilde{\rho}\rho^{k-1} + \cdots + (1-\kappa)^{k-1}\tilde{\rho}^{k-1}\rho)\Big) \\
=&a\Big(-\rho^{k+1} + \frac{\tilde{\rho}\kappa(1-\kappa)^k\tilde{\rho}^k}{1-(1-\kappa)\tilde{\rho}\rho} + \tilde{\rho}\kappa\rho^k(1 + (1-\kappa)\frac{\tilde{\rho}}{\rho} + \cdots + (1-\kappa)^{k-1}\frac{\tilde{\rho}^{k-1}}{\rho^{k-1}})\Big) \\
=&a\Big(-\rho^{k+1} + \frac{\tilde{\rho}\kappa(1-\kappa)^k\tilde{\rho}^k}{1-(1-\kappa)\tilde{\rho}\rho} + \tilde{\rho}\kappa\rho^k \frac{1-\Big((1-\kappa)\frac{\tilde{\rho}}{\rho}\Big)^k}{1-(1-\kappa)\frac{\tilde{\rho}}{\rho}}\Big),
\end{aligned}$$

where the last equality goes through only if $(1-\kappa)\frac{\tilde{\rho}}{\rho} < 1$. Further manipulation of this expression results yields the coefficients reported in the main text: We can conclude that the β from our regressions is:

$$\begin{aligned}
\beta_{k+1} &= \rho^k \frac{1-(1-\kappa)\tilde{\rho}}{1-\tilde{\rho}}\Big(-\rho + \frac{\tilde{\rho}\kappa(1-\kappa)^k\tilde{\rho}^k/\rho^k}{1-(1-\kappa)\tilde{\rho}\rho} + \tilde{\rho}\kappa\frac{1-\left((1-\kappa)\frac{\tilde{\rho}}{\rho}\right)^k}{1-(1-\kappa)\frac{\tilde{\rho}}{\rho}}\Big) \\
&= \rho^k \frac{1-(1-\kappa)\tilde{\rho}}{1-\tilde{\rho}}\left(\frac{\tilde{\rho}-\rho}{1-(1-\kappa)\tilde{\rho}/\rho} + \tilde{\rho}\kappa\left[(1-\kappa)\frac{\tilde{\rho}}{\rho}\right]^k\left(\frac{1}{1-(1-\kappa)\tilde{\rho}\rho} - \frac{1}{1-(1-\kappa)\frac{\tilde{\rho}}{\rho}}\right)\right) \\
&= \rho^k \frac{1-(1-\kappa)\tilde{\rho}}{(1-\tilde{\rho})(1-(1-\kappa)\tilde{\rho}/\rho)}\left((\tilde{\rho}-\rho) + \tilde{\rho}\kappa\left[(1-\kappa)\frac{\tilde{\rho}}{\rho}\right]^k\left(\frac{(1-\kappa)\tilde{\rho}(\rho-\rho^{-1})}{1-(1-\kappa)\tilde{\rho}\rho}\right)\right).
\end{aligned}$$

Under the assumption that $(1-\kappa)\tilde{\rho}/\rho < 1$, the first term is positive and the second one is negative. The second term shrinks with k which implies that the coefficients start negative and the turns positive.

B.3 Proof of Proposition 3

The proof of Proposition 2 reveals that the expression for β_1 is:

$$\beta_1 = \frac{1-(1-\kappa)\tilde{\rho}}{1-\tilde{\rho}}\left(\frac{\tilde{\rho}\kappa}{1-(1-\kappa)\tilde{\rho}\rho} - \rho\right). \tag{A.7}$$

The above expression is negative iff:

$$\begin{aligned}
\tilde{\rho}\kappa &< \rho - (1-\kappa)\tilde{\rho}\rho^2 \\
\tilde{\rho}(\kappa/\rho + (1-\kappa)\rho) &< 1 \\
\tilde{\rho} &< \left[(1-\kappa)\rho + \frac{\kappa}{\rho}\right]^{-1} \equiv \bar{\rho}.
\end{aligned}$$

Note also that $0 < \kappa \leq 1$ implies that $\bar{\rho} \geq \rho$.

B.4 Proof of Proposition 4

Consider once again the expression in (18), which we report here for convenience:

$$\beta_{k+1} = \rho^k \frac{1-(1-\kappa)\tilde{\rho}}{(1-\tilde{\rho})\left(1-(1-\kappa)\frac{\tilde{\rho}}{\rho}\right)}\left\{(\tilde{\rho}-\rho) + \tilde{\rho}\kappa\left[(1-\kappa)\frac{\tilde{\rho}}{\rho}\right]^k\frac{(1-\kappa)\tilde{\rho}(\rho-\rho^{-1})}{1-(1-\kappa)\tilde{\rho}\rho}\right\}. \tag{A.8}$$

Under the assumption that $(1-\kappa)\tilde{\rho}/\rho < 1$, the term outside the curly bracket is positive. Inside the curly bracket, the second term is negative as $\rho - \rho^{-1} < 0$. This term monotonically approaches zero as $k \to \infty$ because we assumed that the term in squared brackets is smaller than 1. The sign of the entire expression then depends on k and on the sign of $\tilde{\rho} - \rho$. If $\tilde{\rho} < \rho$ both signs inside the curly brackets are negative regardless of k, thus $\beta_{k+1} < 0$ for all k. If $0 < \tilde{\rho} < \bar{\rho}$ the first term is positive and the second one is negative. The second term shrinks with k which implies that the coefficients start negative and the turns positive. If $\tilde{\rho} > \bar{\rho}$ then we know from Proposition 3 that $\beta_1 > 0$ which implies that $(\tilde{\rho}-\rho) > \tilde{\rho}\kappa\frac{(1-\kappa)\tilde{\rho}(\rho-\rho^{-1})}{1-(1-\kappa)\tilde{\rho}\rho}$. This last inequality implies that $\beta_{k+1} > 0$ for any k because $(\tilde{\rho}-\rho) > \tilde{\rho}\kappa\left[(1-\kappa)\frac{\tilde{\rho}}{\rho}\right]^k \frac{(1-\kappa)\tilde{\rho}(\rho-\rho^{-1})}{1-(1-\kappa)\tilde{\rho}\rho}$ for any k since $(1-\kappa)\tilde{\rho}/\rho < 1$.

B.5 Proof of Proposition 5

By taking the sum of the β's we obtain:

$$\sum_{k=0}^{\infty}\beta_{k+1} = \sum_{k=0}^{\infty}\rho^k \frac{1-(1-\kappa)\tilde{\rho}}{(1-\tilde{\rho})(1-(1-\kappa)\tilde{\rho}/\rho)}\left\{(\tilde{\rho}-\rho) + \tilde{\rho}\kappa\left[(1-\kappa)\frac{\tilde{\rho}}{\rho}\right]^k\left(\frac{(1-\kappa)\tilde{\rho}(\rho-\rho^{-1})}{1-(1-\kappa)\tilde{\rho}\rho}\right)\right\}$$

$$= \frac{1-(1-\kappa)\tilde{\rho}}{(1-\tilde{\rho})(1-(1-\kappa)\tilde{\rho}/\rho)}\left(\frac{(\tilde{\rho}-\rho)}{1-\rho} + \frac{\tilde{\rho}\kappa}{1-(1-\kappa)\tilde{\rho}}\left(\frac{(1-\kappa)\tilde{\rho}(\rho-\rho^{-1})}{1-(1-\kappa)\tilde{\rho}\rho}\right)\right)$$

So we can conclude our model can explain the excess volatility puzzle as long as:

$$\tilde{\rho} - \rho > \frac{\tilde{\rho}\kappa(1-\rho)}{1-(1-\kappa)\tilde{\rho}}\left(\frac{(1-\kappa)\tilde{\rho}(\rho^{-1}-\rho)}{1-(1-\kappa)\tilde{\rho}\rho}\right) > 0. \tag{A.9}$$

B.6 Proof of Proposition 6

The interest rate process is described in equation (9), while the evolution of interest rate expectations is reported in equations (12)-(13). Using these equations, one can express the one-step ahead forecast errors, $x_t - \mathrm{E}^s_{t-1}\, x_t$, as:

$$x_t - \mathrm{E}^s_{t-1}\, x_t = ((1-\kappa)\tilde{\rho} + \rho)(x_{t-1} - \mathrm{E}^s_{t-2}\, x_{t-1}) - (1-\kappa)\tilde{\rho}\rho(x_{t-2} - \mathrm{E}^s_{t-3}\, x_{t-2}) + \varepsilon_t - \tilde{\rho}\varepsilon_{t-1}$$

Rearranging the above equation, one obtains:

$$x_t - \mathrm{E}^s_{t-1}\, x_t = \frac{(1-\tilde{\rho}L)}{(1-\delta L)(1-\rho L)}\varepsilon_t = \left(\frac{\rho-\tilde{\rho}}{\rho-\delta}\frac{1}{1-\rho L} + \frac{\tilde{\rho}-\delta}{\rho-\delta}\frac{1}{1-\delta L}\right)\varepsilon_t \tag{A.10}$$

where $\delta = (1-\kappa)\tilde{\rho}$. In continuous time, we can express the impulse response of the one-step ahead forecast errors as:

$$m(t) = \frac{\rho-\tilde{\rho}}{\rho-\delta}\rho^t + \frac{\tilde{\rho}-\delta}{\rho-\delta}\delta^t \tag{A.11}$$

where $m(t) = \chi_k$ when $t = k \in \{0, 1, \dots\}$. There is at most one root of $m(t)$, which is:

$$\hat{k} = \frac{\log(\tilde{\rho}-\rho) - \log(\tilde{\rho}-\delta)}{\log\delta - \log\rho} \tag{A.12}$$

Consider first the case in which $\tilde{\rho} \leq \rho$. In such a case, equation (A.10) reveals that the IRF of the forecast errors to an interest rate innovation is always positive.

Consider now the case in which $\tilde{\rho} > \rho$. Equation (A.11) implies that the IRF of the forecast errors is negative for large enough values of t, regardless of whether $\rho - \delta$ is positive or negative. In such a case, a switch in the sign of impulse responses obtains if $\chi_1 = m(1) > 0$.[34] This is the case when $\tilde{\rho} < \hat{\rho} \equiv \frac{\rho}{\kappa}$. Thus, if $\rho < \tilde{\rho} < \hat{\rho}$, then $\chi_k > 0$ for $k < \hat{k}$ and $\chi_k < 0$ for $k > \hat{k}$, where $\hat{k}$ is defined in equation (A.12). Otheriwse, if $\tilde{\rho} \leq \hat{\rho}$, then $\chi_k > 0$ for all $k \geq 1$.

C Alternative Theories of Behavioral Expectations

A common way of modeling behavioral biases is to assume that agents perceive the data generating process to be different from the truth, but then allow agents to update correctly based on their beliefs. We now examine the most popular of such approaches in the context of our model of exchange rate determination.

C.1 Shock Misperception / Adaptive Expectations

Under shock misperception as in GT, agents misperceive the interest rate process to be partially driven by temporary shocks, but have a correct belief about the persistence

[34]Note that $\chi_0 = m(0) = 0$, independently of the parameters governing the process of expectations formation.

parameter $\tilde{\rho} = \rho$. In this case, the coefficients of the excess returns predictability regression take the form:[35]

$$\beta_{k+1} = -[\rho(1-\kappa)]^k \frac{\rho(1-\rho(1-\kappa))(1-\kappa)(1+\rho)}{1-(1-\kappa)\rho^2} < 0. \tag{A.13}$$

For $k = 0$, this expression corresponds to equation (18) in GT. As pointed out in their paper, shock misperception can explain the classic UIP puzzle $\beta_1 < 0$. However, this expression shows that simply relying on misperception about the relative importance of transitory and persistent shocks will not generate a reversal of excess returns (i.e., $\beta_{k+1} < 0$ for all $k \geq 0$) because forecast errors follow an AR(1) process:

$$s - s_t^{RE} = \frac{1}{1-\rho}\Big(\mathrm{E}_t\, x_{t+1} - \mathrm{E}_t^s\, x_{t+1}\Big) = \frac{1}{1-\rho}\Big(\tilde{\rho}(1-\kappa)(\mathrm{E}_{t-1}\, x_t - \mathrm{E}_{t-1}^s\, x_t) + \tilde{\rho}(1-\kappa)\varepsilon_t\Big).$$

The intuition can be understood by thinking about the response of the exchange rate to an interest rate shock. Figure 2a shows such a response for an illustrative case with $\rho = 0.98$ and $\kappa = 0.17$. The Kalman gain and persistent parameters are in the range considered by GT and Ilut (2012).

The red line shows that the exchange rate initially appreciates less than under rational expectations because agents partly believe the shock to be transitory. In the next periods, agents observe the interest rate differential to be higher than what they expected, so they update their beliefs accordingly. That is, over time agents attribute a larger and larger probability to the fact that the observed movements in the interest rate differential are coming from a persistent shock. Eventually, this probability approaches one, thus the exchange rate appreciation will approach the rational expectation appreciation but will never exceed it.

Pure Extrapolation Bias. Under pure extrapolation, agents correctly perceive the interest rate process to be driven exclusively by the persistent component (i.e., $x_t = z_t$) but they have an incorrect perception of the persistence parameter ($\tilde{\rho} \neq \rho$). When $\tilde{\rho} > \rho$, agents are overconfident as to the extent to which current interest rate differentials predict future ones, i.e. $\mathrm{E}_t^s\, x_{t+1} = \tilde{\rho} x_t > \rho x_t = \mathrm{E}_t\, x_{t+1}$. In the Appendix, we show that under pure over-extrapolation the exchange rate predictability coefficients take

[35] This expression obtains as a special case of Proposition 2 when $\tilde{\rho} = \rho$.

the form:

$$\beta_{k+1}^{PEB} = \rho^k \frac{\tilde{\rho} - \rho}{1 - \tilde{\rho}} > 0. \tag{A.14}$$

That is, high interest rate differentials predicts negative excess returns at all horizons. The intuition is simple. As depicted in Figure 2a, following an interest rate increase the exchange rate appreciates more than rational expectations, as interest rates are believed to be more persistent. In the following period, the exchange rate depreciates more than under rational expectations generating negative excess returns. These negative excess returns imply positive β_k at all horizons, as shown in Figure 2b. From equation (A.14), it is easy to see that with pure under-extrapolation ($\tilde{\rho} < \rho$), all the β_k's will be negative so that excess return are positive at all horizons.

Diagnostic Expectations. Consider now a case in which agents know the true data generating process (i.e. $x_t = z_t$, and $\tilde{\rho} = \rho$) but in which a behavioral bias leads agents to forecast the future using a reference conditional distribution that differs from the objective one. In the formulation of Bordalo et al. (2018) and Bordalo et al. (2019), diagnostic expectations about the interest rate process would take the form:

$$\mathrm{E}_t^\theta \, x_{t+1} = \mathrm{E}_t \, x_{t+1} + \theta[\mathrm{E}_t \, x_{t+1} - \mathrm{E}_{t-1} \, x_{t+1}]. \tag{A.15}$$

In essence, diagnostic expectations over-react to the information received at time t by the term $\theta[\mathrm{E}_t \, x_{t+1} - \mathrm{E}_{t-1} \, x_{t+1}]$. It is straightforward to see that this simple version of diagnostic expectations would not lead to the desired result. Suppose that there is an unexpected increase in the interest rate differential at time t. Under diagnostic expectation, s_t appreciates more than under rational expectations. However, absent any other subsequent news, the exchange rate would depreciate at time $t + 1$ to the rational expectation level, and follow the rational expectation path thereafter. The excessive impact appreciation would result in a larger depreciation than under rational expectations in the following period. Thus, on average, an increase in interest rate differential is associated with a subsequent excessive depreciation, the opposite of what the data suggests. Formally:

Lemma 3. *Under diagnostic expectations of the form* (A.15)

$$\beta_1 = \rho(1 + \rho)\theta > 0,$$
$$\beta_{k+1} = 0 \quad \forall k > 0.$$

The second part of the Lemma follows from the fact that expectations in (A.15) react only to current news, so that diagnostic expectations in periods after the shock coincide with rational expectations. A more general formulation of diagnostic expectations discussed in Bordalo et al. (2018, Appendix 2b) would allow over-reaction to past news as well:

$$\mathrm{E}_t^{\theta}(x_{t+1}) = \mathrm{E}_t\, x_{t+1} + \theta \sum_{s\geq 1} \alpha_s \left[\mathrm{E}_{t+1-s}\, x_{t+1} - \mathrm{E}_{t-s}\, x_{t+1}\right]. \tag{A.16}$$

Under the most plausible specification of α's capturing recency effects ($\alpha_1 > \alpha_2 > \ldots$) the coefficients β_k for $k > 1$ would be non-zero but all positive, thus contradicting the empirical reversal.

D Interest Rate Differentials

Here, we estimate the general representation of the interest rate process in (10)-(11) on the *actual* interest rate data. We find strong evidence that for all the countries that the actual interest rate data is best described by the simple AR(1) process in (9), which arises as a special case of (10)-(11) when $\sigma_\nu^2 = 0$.

We estimate a generalized version of (10)-(11) that allows for a constant term:

$$x_t = \mu + z_t + \nu_t, \qquad \nu_t \sim \mathcal{N}(0, \sigma_\nu^2), \tag{A.17}$$

$$z_t = \rho z_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2). \tag{A.18}$$

In the special case of $\sigma_\nu^2 = 0$, the above system collapses to the special case of an AR(1) process. The system (A.17)-(A.18) can be interpreted as a state-space representation, whose parameters $\theta = \{\mu, \rho, \sigma_\varepsilon^2, \sigma_\nu^2\}$ can be estimated via Maximum Likelihood.

We estimate the system using end-of-the-quarter observations on the 3-month Euro-market rates for Canada, Germany, Japan, Switzerland, and United Kingdom against the the 3-month Euro-dollar interest rate. Our sample period ranges from 1986:M8 to 2007:M12.[36] The results from the estimation are reported in Table A.1.

[36] We choose a quarterly frequency for coherence with Section 4, which uses the data on interest rate expectations at the shortest forecast horizon available, i.e., 3-months. Nevertheless, in both sections we use all the of information available in the monthly data by using all three quarterly time series that can be constructed for each country from monthly data: the ones constructed from the first, second, and third month of each quarter, respectively.

Table A.1: ML estimates of the state-space system on interest rates data

	Canada	Germany	UK	Japan	Switzerland
μ	-0.007 (0.008)	0.006 (0.008)	-0.022 (0.005)	0.028 (0.007)	0.018 (0.005)
ρ	0.918 (0.006)	0.963 (0.013)	0.922 (0.012)	0.959 (0.014)	0.952 (0.017)
$100\sigma_\varepsilon$	0.596 (0.053)	0.554 (0.025)	0.696 (0.044)	0.566 (0.025)	0.624 (0.027)
$100\sigma_\nu$	0.110 (0.136)	0.000 (0.037)	0.082 (0.173)	0.000 (0.043)	0.000 (0.051)
$\sigma_\nu = 0$ (p-value)	0.686	1.000	0.820	1.000	1.000

The table shows that the interest rate differential is a very persistent process for all the economies considered, with the estimates of ρ ranging from 0.918 (Canada) to 0.963 (Germany). The table also shows that in all cases the null hypothesis of $\sigma_\nu = 0$ cannot be rejected at any significance level for any country. These findings are in line with the existing literature (e.g, Gourinchas and Tornell (2004); Ilut (2012)) and lend support to the assumption we made in the model that the true interest rate differential is best characterized as a very persistent autoregressive process.

E Derivations of GE model

E.1 Households' optimality conditions

The households' maximization problem yields the following optimality conditions:

$$C_t^\sigma L_t^\varphi = \frac{W_t}{P_t}, \tag{A.19}$$

$$1 \geq \mathbb{E}_t\left[\Theta_{t,t+1}\right] + \frac{D_t^{-\iota}}{C_t^{-\sigma}}, \tag{A.20}$$

$$1 \geq \mathbb{E}_t\left[\Theta_{t,t+1} R_t\right] + \eta\frac{D_t^{-\iota}}{C_t^{-\sigma}}, \tag{A.21}$$

$$1 = \mathbb{E}_t\left[\Theta_{t,t+1} R_t^m\right], \tag{A.22}$$

$$1 = \mathbb{E}_t\left[\Theta_{t,t+1} R_t^{m*} \frac{S_{t+1}}{S_t}\right] - \delta\left(\frac{S_t B_{H,t+1}^{m*}}{P_t} - b^{m*}\right). \tag{A.23}$$

where we have defined the nominal stochastic discount factor of the home household as: $\Theta_{t,t+s} = \beta\frac{\zeta_{t+s}}{\zeta_t}\left(\frac{C_{t+s}}{C_t}\right)^{-\sigma}\frac{P_t}{P_{t+s}}$.

E.2 Steady state

In this sections, we derive the log-linear UIP condition of the model in Section 4. We start by characterizing the steady state of the model. We consider the symmetric steady state of the model where $A = A* = \zeta = \zeta^* = 1$ and $B_H^m = B_F^m = B_H^{m*} = B_F^{m*} = 0$. In the symmetric steady state, exchange rates and the terms of trade are equal to 1: $S = \mathcal{Q} = \mathcal{T} = 1$. We normalize prices such that $P_H = P_F^* = 1$, which imply $P_t = P_t^* = 1$. Evaluating aggregate demand at the steady state we obtain:

$$Y_H = (1-\gamma)C + \gamma C_t^*,$$
$$Y_F = \gamma C + (1-\gamma)C_t^*,$$

which imply $Y_H = Y_F = C_t = C_t^*$. The marginal cost in steady state is given by $\frac{W}{P}\frac{1}{A} = \frac{\nu-1}{\nu}$. Combining these expressions with the labor supply condition we obtain:

$$\chi Y_H^{\sigma+\psi} = \frac{\nu-1}{\nu}.$$

We set $\chi = \frac{\nu-1}{\nu}$ so that $Y_H = L = 1$. By symmetry $Y_F = L^* = 1$. Next, we evaluate the intertemporal conditions for the household:

$$1 \geq \beta + D^{-\iota}, \tag{A.24}$$
$$1 \geq \beta R + \eta D^{-\iota}, \tag{A.25}$$
$$1 = \beta R^m, \tag{A.26}$$
$$1 = \beta R^{m*}, \tag{A.27}$$

where we have used the fact that $\Theta = \beta/\Pi$ and $\Pi = 1$. We consider a steady state with a positive supply of government bonds financed by lump-sum taxes so that $\frac{B}{P} = b > 0$. Market-clearing condition requires $b_H = b > 0$, which implies that condition (A.25) holds with equality. Substituting the definition of D in that equation we obtain:

$$1 = \beta R + \eta(m - \eta b)^{-\iota} \tag{A.28}$$

where $m = M_{H,t}/P$ and where we have imposed the market clearing condition in the market for home money. This equation makes clear that the monetary authority controls the interest rate by choosing the money supply or by engaging in open-market operations of government bonds in exchange for money (given that $\eta < 1$). It remains to verify that the constraint $M = m > 0$ which is true if and only if (A.24) holds with equality. We proceed by guessing that the latter is true and verify that indeed $m > 0$:

$$1 = \beta + (m - \eta b)^{-\iota} \implies m = \eta b + (1-\beta)^{-\iota} > 0.$$

A similar argument applies to the foreign equivalent of (A.24)-(A.25). When we solve the dynamic model using a first-order approximation around this steady state, we assume that shocks are sufficiently small that these conditions keep holding with equality.

E.3 Derivation of UIP condition in GE model

In this section, we derive the log-linear UIP condition of the model in Section 4. The home household's relevant optimality conditions are equations (A.20)-(A.23), and the foreign household's counterparts are:

$$1 = \mathbb{E}_t \left[\Theta^*_{t,t+1}\right] + \frac{(D^*_t)^{-\iota}}{C_t^{*-\sigma}}, \tag{A.29}$$

$$1 = \mathbb{E}_t \left[\Theta^*_{t,t+1} R^*_t\right] + \eta \frac{(D^*_t)^{-\iota}}{C_t^{*-\sigma}}, \tag{A.30}$$

$$1 = \mathbb{E}_t \left[\Theta^*_{t,t+1} R^{m*}_t\right], \tag{A.31}$$

$$1 = \mathbb{E}_t \left[\Theta^*_{t,t+1} R^m_t \frac{S_t}{S_{t+1}}\right] - \delta \left(\frac{B^m_{F,t+1}}{S_t P^*_t} - b^m\right). \tag{A.32}$$

where we have defined $\Theta_{t+1} = \beta \frac{\zeta^*_{t+1}}{\zeta^*_t} \left(\frac{C^*_{t+1}}{C^*_t}\right)^{-\sigma} \frac{P^*_t}{P^*_{t+1}}$. An implication of the no-arbitrage conditions (A.22), (A.23), (A.31), (A.32) is that an increase in real foreign asset holding by the home country should be compensated by a decrease in the foreign country:

$$\delta \left(\frac{S_t B^{m*}_{H,t+1}}{P_t} - b^{m*}\right) = -\delta \left(\frac{B^m_{F,t+1}}{S_t P^*_t} - b^m\right).$$

Now define $\tilde{D}_t = \frac{D_t^{-\iota}}{C_t^{-\sigma}}$ and $\tilde{D}^*_t = \frac{D_t^{*-\iota}}{C_t^{*-\sigma}}$. The log-linear version of equations (A.20)-(A.23) and (A.29)-(A.32) read:

$$0 = \mathrm{E}_t\,\theta_{t+1} + \frac{1-\beta}{\beta}(\tilde{d}_t), \tag{A.33}$$

$$0 = \mathrm{E}_t\,\theta_{t+1} + r_t + \frac{(1-\beta)\eta}{1-(1-\beta)\eta}(\tilde{d}_t), \tag{A.34}$$

$$0 = \mathrm{E}_t\,\theta_{t+1} + r_t^m, \tag{A.35}$$

$$0 = \mathrm{E}_t\,\theta_{t+1} + r_t^{m*} + \mathrm{E}_t\,\Delta s_{t+1} - \delta(b_t^{m*}), \tag{A.36}$$

$$0 = \mathrm{E}_t\,\theta_{t+1}^* + \frac{1-\beta}{\beta}(\tilde{d}_t^*), \tag{A.37}$$

$$0 = \mathrm{E}_t\,\theta_{t+1}^* + r_t^* + \frac{(1-\beta)\eta}{1-(1-\beta)\eta}(\tilde{d}_t^*), \tag{A.38}$$

$$0 = \mathrm{E}_t\,\theta_{t+1}^* + r_t^{m*}, \tag{A.39}$$

$$0 = \mathrm{E}_t\,\theta_{t+1}^* + r_t^m - \mathrm{E}_t\,\Delta s_{t+1} - \delta(b_t^m). \tag{A.40}$$

where we have defined $b_t^m = \frac{B_{F,t+1}^m}{S_t P_t^*}$ and $b_t^{m*} = \frac{S_t B_{H,t+1}^{m*}}{P_t}$. Combine equations (A.33)-(A.34) (and the analogous foreign economy equations (A.37)-(A.38)) to obtain:

$$0 = \mathrm{E}_t\,\theta_{t+1} + r_t + \alpha r_t, \tag{A.41}$$

$$0 = \mathrm{E}_t\,\theta_{t+1}^* + r_t^* + \alpha r_t^*, \tag{A.42}$$

where $\alpha = \beta\frac{\eta}{1-\eta} > 0$. Combine equations (A.35)-(A.36) (or the analogous foreign economy equations (A.39)-(A.40)) to obtain:

$$\mathrm{E}_t\,\Delta s_{t+1} - (r_t^m - r_t^{m*}) = \delta b_t^m. \tag{A.43}$$

One can use equations (A.35), (A.39), (A.41), (A.42), and (A.43) to show that:

$$\mathrm{E}_t\,\Delta s_{t+1} - (r_t - r_t^*) = \alpha(r_t - r_t^*) + \delta b_t^m,$$

which is equation (32) in the text.

E.4 Interest rate dynamics under the PM/AF regime

In this section, we show that an upward revision of today's interest rate is systematically associated with a downward revision of the cumulative path of interest rates when tax revenues are constant in a PM/AF regime.

Consider the log-linear version of a special case of the government budget constraint (28) that obtains when $T_t = 0$ for all t:

$$R^{-1}(b_{H,t+1} - r_t) = b_{H,t} - \pi_t$$

Solving this budget constraint forward and applying the No-Ponzi-Game Condition, $\lim_{k\to\infty} R^{-k} b_{H,t+k} = 0$, we obtain:

$$b_{H,t} = (1 - R^{-1}) \sum_{k=0}^{\infty} R^{-k} \pi_{t+k} = \frac{1 - \phi_\pi R^{-1}}{\phi_\pi} \sum_{k=0}^{\infty} R^{-k} r_{t+k}, \tag{A.44}$$

where we used $r_t = \phi_\pi \pi_t, \ \forall t$. Equation (A.44) implies that the path of nominal interest rates needs to be consistent with the current stock of debt at every point in time. In particular, the discounted sum of nominal interest rates has to be proportional to the current stock of debt to ensure that government debt is solvent. In addition, equation (A.44) places restrictions on the properties of agents' revisions of the interest rate path. That is:

$$\mathrm{E}_t\, b_{H,t} - \mathrm{E}_{t-1}\, b_{H,t} = \frac{1 - \phi_\pi R^{-1}}{\phi_\pi} \sum_{k=0}^{\infty} R^{-k} \left(\mathrm{E}_t\, r_{t+k} - \mathrm{E}_{t-1}\, r_{t+k} \right), \tag{A.45}$$

implying that any revision of the interest rate path must satisfy the government budget constraint. Note that, because $b_{H,t}$ is known at $t-1$, $\mathrm{E}_t\, b_{H,t} - \mathrm{E}_{t-1}\, b_{H,t} = 0, \ \forall t$, and equation (A.45) becomes:

$$\sum_{k=0}^{\infty} R^{-k} \left(\mathrm{E}_t\, r_{t+k} - \mathrm{E}_{t-1}\, r_{t+k} \right) = 0. \tag{A.46}$$

Thus, the government budget constraint implies that the discounted sum of interest rate revisions is always equal to zero. In other words, any revision in today's interest rate must be associated with a revision of future interest rates that leaves the discounted sum unchanged. If we allow the path of interest rate revisions to change sign at most once, an upward revision in today's interest rate requires a downward revision of interest rates from $t + \hat{k}$ onward so that the discounted sum of interest rates is unchanged, where $\hat{k}$ is the time of the switch in sign. In addition, because equation (A.46) features discounting, revisions of future interest rates must be larger, in absolute value, than revisions of current interest rates. It follows that an upward revision in today's interest rate implies a downward revision of the *undiscounted* sum of interest rates. As a result:

$$\text{Cov}\left(\mathrm{E}_t\, r_t - \mathrm{E}_{t-1}\, r_t, \sum_{k=0}^{\infty}\left(\mathrm{E}_t\, r_{t+k} - \mathrm{E}_{t-1}\, r_{t+k}\right)\right) < 0. \tag{A.47}$$

This covariance naturally applies also to interest rate differentials, and it has the opposite sign relative to the baseline AM/PF regime (where interest rates feature monotonic responses). Because exchange rate changes reflect, at least in part, revisions of the undiscounted sum of interest rate differentials (see eq. (15)) understanding dynamic properties of this object is important for exchange rate dynamics. As we argue in Section 4.3, the result in equation (A.47) implies that systematic under-reaction to interest rate increases generates short-run depreciations of exchange rates in excess of UIP.

E.5 Solution method

In this section, we describe how to solve a dynamic general equilibrium linear model with distorted beliefs. The part of the solution that involves solving a rational expectations model under incomplete information draws on Blanchard et al. (2013). The part of the solution that deals with the belief distortions is our own algorithm.

The solution proceeds in two steps. First, we write down the representation of the model perceived by the agents. Crucially, because agents expectations are rational conditional on their beliefs about the model, we can solve for agents expectations using standard techniques. In a second step, we link agents' expectation errors to the true structural shocks. Because agents have an incorrect perception of the exogenous processes, agents expectation errors will turn out to be autocorrelated. These autocorrelated expectation errors are the source of systematic, and thus predictable, mistakes that agents make in the model.

Recall that the exogenous processes $x_t \in \{\log A_t, \log A_t^*, \log \zeta_t, \log \zeta_t^*\}$ are perceived to follow:

$$\begin{aligned} x_t &= z_t^x + \nu_t^x, & \nu_t^x &\sim \mathcal{N}(0, \tilde{\sigma}_{\nu,x}^2), \\ z_t^x &= \tilde{\rho} z_{t-1}^x + \varepsilon_t^x, & \varepsilon_t^x, &\sim \mathcal{N}(0, \sigma_x^2). \end{aligned}$$

Agents observe x_t but do not observe separately z_t^x and ν_t^x. The unobserved state

$\mathbf{z}_t = (z_t^A, z_t^{A^*}, z_t^\zeta, z_t^{\zeta^*})'$ follows the process:

$$\mathbf{z}_t = \tilde{\mathbf{A}}\mathbf{z}_{t-1} + \mathbf{B}\varepsilon_t, \tag{A.48}$$

where $\varepsilon_t = (\varepsilon_t^A, \varepsilon_t^{A^*}, \varepsilon_t^\zeta, \varepsilon_t^{\zeta^*})'$, the matrix $\mathbf{A} = \tilde{\rho}\mathbf{I}_4$, and the vector $B = (1,1,1,1)'$. The representative agents observes the vector $\mathbf{s}_t = (x_t^A, x_t^{A^*}, x_t^\zeta, x_t^{\zeta^*})'$:

$$\mathbf{s}_t = \mathbf{C}\mathbf{z}_t + \mathbf{D}\nu_t, \tag{A.49}$$

where the vector $\nu_t = (\nu_t^A, \nu_t^{A^*}, \nu_t^\zeta, \nu_t^{\zeta^*})'$, the matrix $\mathbf{C} = \mathbf{I}_4$, and the vector $D = (1,1,1,1)'$. Let $\mathbf{y}_t$ denote the vector of endogenous variables, which includes endogenous states. The equilibrium condition of the model can be described by the following difference equation:

$$\mathbf{F}\,\mathrm{E}_t[\mathbf{y}_t] + \mathbf{G}\mathbf{y}_t + \mathbf{H}\mathbf{y}_{t-1} + \mathbf{M}\mathbf{s}_t + \mathbf{N}\,\mathrm{E}_t[\mathbf{s}_{t+1}],$$

where $\mathbf{F}, \mathbf{G}, \mathbf{H}, \mathbf{M}, \mathbf{N}$ are matrices of parameters. Notice that the unobservable exogenous state $\mathbf{z}_t$ only enters the equilibrium through the observable vector $\mathbf{s}_t$, reflecting the assumption that the information set of the representative agent is given only by past and current values of $\mathbf{s}_t$ and of the endogenous state $\mathbf{y}_t$. Suppose there is a unique stable solution of the model that takes the form:

$$\mathbf{y}_t = \mathbf{P}\mathbf{y}_{t-1} + \mathbf{Q}\mathbf{s}_t + \mathbf{R}\mathbf{x}_{t|t}, \tag{A.50}$$

where $\mathbf{x}_{t|t}$ denotes agents expectations $E[\mathbf{x}_t|\mathbf{s}_t, \mathbf{s}_{t-1}, \ldots]$. The matrices $\mathbf{P}, \mathbf{Q}, \mathbf{R}$ can be found by solving the following matrix equations:

$$\begin{aligned} \mathbf{F}\mathbf{P}^2 + \mathbf{G}\mathbf{P} + \mathbf{H} &= \mathbf{0}, \\ (\mathbf{F}\mathbf{P} + \mathbf{G})\mathbf{Q} + \mathbf{M} &= \mathbf{0}, \\ (\mathbf{F}\mathbf{P} + \mathbf{G})\mathbf{R} + [\mathbf{F}(\mathbf{Q}\mathbf{C} + \mathbf{R}) + \mathbf{N}\mathbf{C}]\mathbf{A} &= \mathbf{0}. \end{aligned}$$

The matrix $\mathbf{P}$ can be solved for using the technique described in Uhlig et al. (1995). The solution of the other two matrices $\mathbf{Q}$ and $\mathbf{R}$ is straightforward as they enter the system linearly.

We can now use the Kalman filter to describe the evolution of agents' expectations as

$$\mathbf{x}_{t|t} = \tilde{\mathbf{A}}\mathbf{x_{t-1|t-1}} + \mathbf{K}(\mathbf{s_t} - \mathbf{s_{t|t-1}}) \tag{A.51}$$

where $\mathbf{K} = \kappa\mathbf{I}_4$ is the matrix of Kalman gains associated with the signal extraction problem in (A.48)-(A.49). Recall that the true exogenous processes follow AR(1) processes with true persistence parameter ρ so they can be written compactly as:

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\varepsilon_t,$$

where $\mathbf{A} = \rho\mathbf{I}_4$. The last step of the solution involves mapping the signal vector $\mathbf{s_t}$ and the signal forecast errors $\mathbf{s_t} - \mathbf{s_{t|t-1}}$ to the *true* exogenous processes and structural shocks:

$$\mathbf{s}_t = \mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\varepsilon_t \tag{A.52}$$

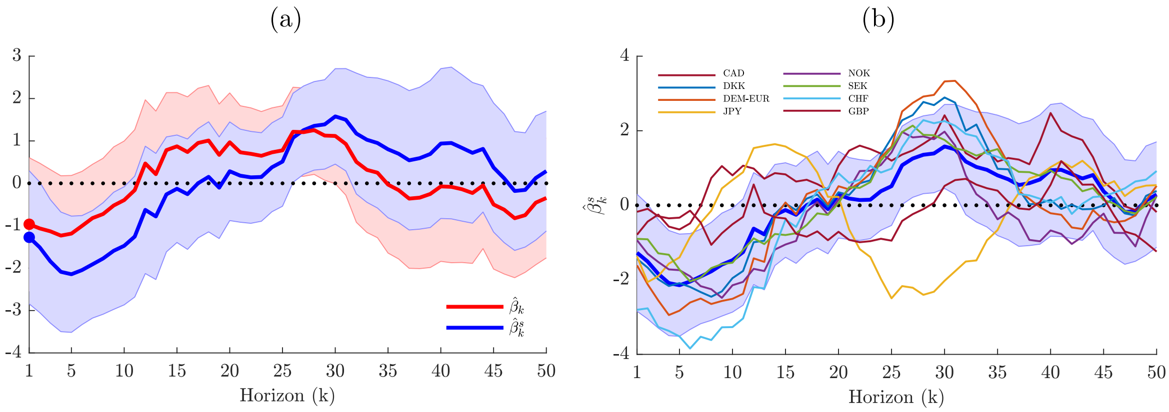
$$\mathbf{s_t} - \mathbf{s_{t|t-1}} = \Phi_1(\mathbf{s_{t-1}} - \mathbf{s_{t-1|t-2}}) + \Phi_2(\mathbf{s_{t-2}} - \mathbf{s_{t-2|t-3}}) + \mathbf{B}\varepsilon_t - \tilde{\rho}\mathbf{B}\varepsilon_{t-1} \tag{A.53}$$

where we have defined $\Phi_1 = ((1-\kappa)\tilde{\rho} + \rho)\mathbf{I}_4$, $\Phi_2 = -(1-\kappa)\tilde{\rho}\rho\mathbf{I}_4$. The evolution of the signal forecast error in (A.53) follows directly from the proof of Proposition 6.

To summarize the solution of the model, we first solve for the expectations of the agents under the perceived model and obtain (A.50), which is the the law of motion of the endogenous variables in terms of the agents' signals, $\mathbf{s}_t$ and the agents' current belief of the state, $\mathbf{x}_{t|t}$. Then, we express $\mathbf{s}_t$ and $\mathbf{x}_{t|t}$ in terms of the *true* structural shocks using the *true* mapping (A.52)-(A.53) and the Kalman filter equation (A.51). Because (A.50) is the rational expectations solution conditional on the perceived model, the policy rules in (A.50) reflect agents' beliefs that the innovations in their signals are *i.i.d.*. Nevertheless, by then plugging in (A.53) it becomes evident that these signals' innovations are truly autocorrelated, reflecting the systematic mistakes that agents make under distorted beliefs.

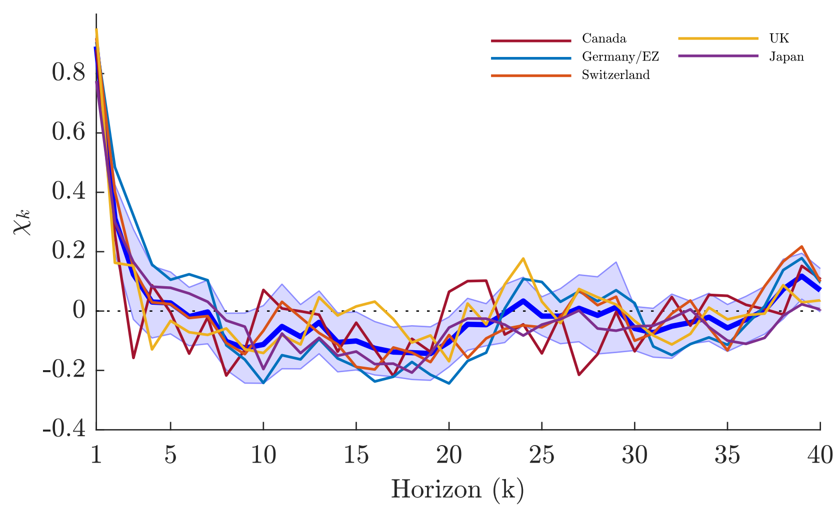
F Additional Figures

Figure A.1: UIP Regression Results (1990:M1-2019:M7)



Notes: Panel (A.1a) reports the panel estimates of β_k and β^s_k in equations (3) and (7), respectively, along with their 95% confidence interval. Panel (A.1b) reports the country-by-country estimates of β^s_k, along with the panel counterpart. Each horizon is a quarter.

Figure A.2: IRFs of the Forecast Errors in the Interest Rate Differentials (1986:M8-2020:M6)



Notes: The figure reports the empirical impulse response of the one-quarter ahead forecast error in the interest rate differential to an innovation in the interest rate differential (see Eq. (22)). Each horizon is a quarter.