Ambiguity Aversion: Implications for the Uncovered Interest Rate Parity Puzzle

By Cosmin Ilut*

High-interest-rate currencies tend to appreciate in the future relative to low-interest-rate currencies instead of depreciating as uncovered-interest-parity (UIP) predicts. I construct a model of exchange-rate determination in which ambiguity-averse agents face a dynamic filtering problem featuring signals of uncertain precision. Solving a max-min problem, agents act upon a worst-case signal precision and systematically underestimate the hidden state that controls payoffs. Thus, on average, agents next periods perceive positive innovations, which generates an upward re-evaluation of the strategy's profitability and implies ex-post departures from UIP. The model also produces predictable expectational errors, negative skewness and time-series momentum for currency speculation payoffs.
JEL: F31, G11, G14, D80
Keywords: uncovered interest rate parity, carry trade, ambiguity aversion, robust filtering, crash risk, momentum

According to uncovered interest rate parity (UIP), periods when the domestic interest rate is higher than the foreign interest rate should on average be followed by periods of domestic currency depreciation. An implication of UIP is that a regression of realized exchange rate changes on interest rate differentials should produce a coefficient of one. This implication is strongly counterfactual. In practice, UIP regressions (Hansen and Hodrick (1980), Fama (1984)) produce coefficient estimates well below one and sometimes even negative.[1] This anomaly, referred to as the UIP puzzle, is taken very seriously because the UIP equation is a property of most open economy models.

My explanation for the puzzle assumes that investors are ambiguity averse. Attempts to exploit interest rate differentials typically involve borrowing in the low interest rate currency and lending in the high interest rate currency, a strategy known as the 'carry trade'. Such a strategy exposes investors to currency rate risk, particularly the risk that the high interest rate currency will depreciate.

* Duke University, Department of Economics, 213 Social Sciences Bldg., Durham, NC, 27708, E-mail: cosmin.ilut@duke.edu. I would like to thank Gadi Barlevy, Craig Burnside, Larry Christiano, Eddie Dekel, Martin Eichenbaum, Lars Hansen, Peter Kondor, Giorgio Primiceri, Sergio Rebelo, Tom Sargent, Martin Schneider, Tomasz Strzalecki, Eric van Wincoop, Moto Yogo and seminar participants at the NBER IFM Meeting 2010, AEA Annual Meeting 2010, MNB 2009, SED 2009, SITE, 2009 and Board of Governors, Chicago Fed, Duke, ECB, Minneapolis Fed, New York Fed, NYU, Northwestern, Philadelphia Fed, UC Davis, UC Santa Cruz, Univ. of Virginia for helpful discussions and comments.

[1] There is a very large empirical literature on documenting the UIP puzzle. Among recent studies see Gourinchas and Tornell (2004), Burnside et al. (2011), Verdelhan (2010) and Engel (2011).

An ambiguity averse investor rationally places greater weight on such negative outcomes. This reduces the perceived profitability of the investment strategy relative to the true data generating process, thereby leaving some of the gains unexploited. Thus, the model presented here attributes the ex-post failure of UIP and the average positive ex-post profits of the carry trade strategy to the ex-ante uncertainty faced by an ambiguity averse investor.

In this paper, I analyze an environment where rational agents are not endowed with complete knowledge of the true data generating process (DGP) and where they confront this uncertainty with ambiguity aversion. I model ambiguity aversion along the lines of the max-min expected utility (or multiple priors) preferences as axiomatized in Gilboa and Schmeidler (1989). In endogenizing the exchange rate determination I take as exogenous the domestic/foreign interest rate differential. The latter is described as a stochastic process, in which the observed differential is the sum of unobserved persistent and transitory components. I assume that the ambiguity averse investor is uncertain about the variances of the innovations in the temporary and persistent components and that she allows for the possibility that those variances change over time. In other words, the agent perceives the signals she receives about the hidden persistent state as having uncertain precision.

The ambiguity averse agent behaves *as if* she maximizes expected utility under a worst-case belief that is chosen from a set of conditional probabilities. The agent decides how many foreign currency bonds to buy and sell subject to the worst-case belief. The belief is chosen so that, conditional on the agent's bond decision, expected utility is minimized subject to a particular constraint. The constraint is that the agent only considers an exogenously-specified finite set of values for the unknown variances. I choose this set so that the variance parameters selected by the agent are not implausible in a likelihood ratio sense.

The intuition for the model's ability to explain the ex-post failure of UIP starts from understanding the equilibrium reasoning behind an ambiguity averse agent's investment decision. In equilibrium, the agent invests in the higher interest rate bond (investment currency) by borrowing in the lower interest rate bond (funding currency). Faced with uncertainty, the agent chooses to act on beliefs that, compared to the true DGP, underestimate the hidden state of the investment differential, i.e. the differential between the higher interest rate and the lower interest rate. Given the assumed structure of uncertainty, the agent underestimates the hidden state by reacting in equilibrium asymmetrically to signals about it. The agent acts *as if* it is more likely that observed increases in the investment differential have been generated by temporary shocks (a low precision of signals), while observed decreases by persistent shocks (a high precision of signals). The UIP condition holds ex-ante under these endogenously pessimistic beliefs.

The systematic underestimation of the investment differential's hidden state implies that the agent perceives on average positive innovations when updating the estimate. This higher estimate raises the present value of the future payoffs of investing in the higher interest rate bond. The updating effect leads to a higher demand for the investment currency which drives up further its value. Thus, an investment currency could see a subsequent ex-post equilibrium appreciation instead of a depreciation as UIP predicts. This is a manifestation of the ex-post failure of UIP.

Simulation results show that the benchmark specification generates an asymptotically negative UIP regression coefficient. In small samples the magnitude of the coefficient is similar but it is less significant statistically. The benchmark specification also imposes constraints on the set of possible distortions implying that the equilibrium distorted probability distribution is statistically close to the true one. If these constraints are relaxed, the same qualitative results hold but quantitatively they become stronger.

The solution to the UIP puzzle presented in this model produces a unified explanation for three well-documented empirical characteristics of currency market dynamics: ex-post failure of UIP, asymmetric exchange rate behavior and time-series momentum. The asymmetric exchange rate behavior refers to the observation that high-interest-rate currencies tend to appreciate slowly but depreciate suddenly. This asymmetry, recently described as "crash risk", is manifested in the negative skewness of carry trade payoffs, a property documented for example in Brunnermeier, Nagel and Pedersen (2008) and Burnside et al. (2011). My model generates the negative skewness as a result of the asymmetric response to news: in equilibrium the agent underreacts to increases in the high interest rate leading to slower appreciations of the investment currency. However, the agent also overreacts to decreases in the high interest rate which produce sudden depreciations of the investment currency.

The time-series momentum refers to the positive predictability of a currency excess payoff based on its own past payoff.[2] One manifestation of momentum is the so-called 'delayed overshooting puzzle'. This is an empirically documented impulse response in which following a positive shock to the domestic interest rate, the domestic currency experiences a gradual appreciation for several periods instead of an immediate appreciation and then a path of depreciation as UIP implies.[3] The reason behind the model's ability to generate momentum and the delayed overshooting puzzle is the equilibrium underreaction to 'good news'. As in the rational expectations model, a positive innovation in the investment differential causes the investment currency to appreciate. However, the gradual incorporation of this positive innovation predicts time-series momentum: an econometrician analyzing the data generated by the model observes further average appreciations of the investment currency.

The remainder of the paper is organized as follows. Section I relates the pa-

[2]Jorda and Taylor (2009) and Moskowitz, Ooi and Pedersen (2011) find that currencies that had recent positive excess returns continue to do so in the near future. Burnside, Eichenbaum and Rebelo (2011) find that traditional risk factors do not explain the ex-post profitability of momentum strategies.

[3]See Eichenbaum and Evans (1995), Faust and Rogers (2003) and Scholl and Uhlig (2008) for such empirical evidence.

per with the existing literature. Section II describes and discusses the model. Section III presents the model's solution and its implications for exchange rate determination. Section IV describes the results and Section V concludes.

I. Relation to the literature

The standard approach in addressing the UIP puzzle has been to assume a representative agent model with rational expectations and time-varying risk premia. This approach has been criticized in two ways: first, survey evidence has been used to cast doubt on the rational expectations assumption. In this type of work, as for example in Frankel and Froot (1989) and Bacchetta, Mertens and van Wincoop (2009), survey data has been used to argue that most of the predictability of currency excess returns is due to expectational errors.[4]

Second, other empirical research challenges the risk implications of the analysis. The earlier literature, surveyed in Engel (1996), rejected the risk explanation based on the observation that a representative agent model with standard utility functions cannot generate these large and variable risk premia. Given this difficulty, a more recent strand of literature has investigated non-standard utility functions combined with assumptions about the time-variation of the data generating process. Some of this work includes Farhi and Gabaix (2008), Verdelhan (2010), and Colacito and Croce (2011). There is also a recent empirical literature that argues that there is some evidence for risk-premia in currency markets, as in Lustig and Verdelhan (2007) and Lustig, Roussanov and Verdelhan (2011). Burnside et al. (2011) and Engel (2011) offer a critical review of recent risk-based explanations.

The model presented here belongs to an alternative approach, which attempts to explain the failure of UIP through expectational errors. In the existing models, expectational errors arise because of broadly defined behavioral biases or due to small sample Bayesian learning. Examples of models with investors characterized by behavioral biases are Mark and Wu (1998), Gourinchas and Tornell (2004) and Burnside et al. (2010). Froot and Thaler (1990) informally argue that if some agents are slow to respond to news this may explain the puzzle. Bacchetta and van Wincoop (2010) formalize this idea by assuming infrequent portfolio adjustment in the presence of costly currency management. Lewis (1989) argues that Bayesian learning about infrequent shifts in the distribution may be able to explain the UIP puzzle in a small sample.

The closest related paper out of this literature is Gourinchas and Tornell (2004). There they show how an ad-hoc, time-invariant, systematic underreaction to signals about the time-varying hidden state of the interest rate differential can explain the UIP and the delayed overshooting puzzle. Thus, the reduced form

[4]More specifically, these papers regress the survey expectations of future relative currency movements on the observed interest rate differentials. In these regressions, the UIP coefficient is much closer to one, which suggests that the UIP equation holds much better under the survey expectations than under the rational expectations assumption.

mechanism through which my model explains the UIP puzzle is similar. The key difference is that I investigate a model which addresses the origin and optimality of such beliefs. This optimality also implies that the rational agent in this model acts cautiously and underreacts only to good news while overreacting to bad news about her investment strategy.

There are several essential features that distinguish this model from the literature on expectational errors. First, the agents in this model are rational. Here the expectational errors arise when the ambiguity averse agent facing Knightian uncertainty is rationally weighing more the less favorable possible DGPs, which might differ from the true DGP. Second, in this model the response to news is time-varying to reflect the endogenous worst-case signal evaluation of the agent. The endogeneity here refers to the investment position: irrespective of the model's details for the exogenous process, the agent involved in the carry trade is concerned about the future depreciation of the investment currency. This argument highlights that the model is likely to be robust to the specific stochastic process assumptions, which is in stark contrast with the exogenously assumed reactions in behavioral-based models. Third, the model has the ability to produce a unified explanation for the UIP puzzle, time-series momentum and the "crash risk" based on the same structural mechanism.[5] Fourth, in contrast to models based on small sample effects, ambiguity aversion is a preference that naturally ensures that uncertainty about the true DGP can persist even in the limit. In fact, some of the model's results are even stronger in large samples.

The model also contributes more generally to the theoretical literature on learning under ambiguity and asset pricing.[6] The structure of uncertainty that I investigate, namely signals of uncertain precision, is similar to Epstein and Schneider (2007, 2008). There they essentially analyze learning about a constant hidden parameter for which the direction of the worst-case scenario is constant. There are two main modeling differences here. The first is that I consider a setup where the direction of the worst-case interpretation of signals is time-varying. This interpretation is driven by the investment position, which here switches from investing in one currency or the other, depending on where the interest rate is higher. The second, and more important, difference is that in this model the persistent hidden state is time-varying so that the ambiguity-averse agent solves a dynamic filtering problem. The model described in this paper can be interpreted as a general environment in which a dynamic filtering problem with ambiguous signals can be used to explain time-varying predictability and negative skewness of ex-post excess returns through systematic expectational errors.

[5]Some models take as given the "crash risk" and aim to explain the UIP puzzle through this risk (such as Farhi and Gabaix (2008)). Plantin and Shin (2011) take as given the UIP puzzle and generate endogenously the asymmetric exchange rate behavior.

[6]See, among others, Epstein and Schneider (2007), Hansen (2007), Hansen and Sargent (2007), Epstein and Schneider (2008), Leippold, Trojani and Vanini (2008) and Hansen and Sargent (2010).

II. Model

A. Basic Setup

The key model ingredients are: an exogenous interest rate differential process that is the sum of unobserved persistent and transitory components and ambiguity averse investors that are uncertain about the variance of the temporary component. Besides ambiguity, the setup is an otherwise typical one good, two-country, dynamic model of endogenous exchange rate determination similar to a simplified version of Bacchetta and van Wincoop (2010). I now describe the main elements of the model: the assets, the agents, the investment opportunities, the market clearing condition and finally the optimization problem of the investors.

Assets. — The description of the nominal assets available follows closely Bacchetta and van Wincoop (2010). The model contains one good and three assets. There is one good for which purchasing power parity (PPP) holds: $p_t = p_t^* + s_t$, where p_t is the log of price level of the good in the Home country and s_t is the log of the nominal exchange rate defined as Home currency (USD) units per one unit of Foreign currency. Foreign country variables are indicated with a star. The three assets are: one-period nominal bonds in both currencies issued by the respective governments and a risk-free technology with real return $\overline{r}$. Domestic and Foreign bonds are in fixed supply in the Home and Foreign currency respectively.

The Home and Foreign nominal interest rates are i_t and i_t^* respectively. The Home country follows a monetary policy that commits to a constant price level, normalized to one. This implies zero Home inflation so that by no-arbitrage the Home nominal interest rate $i_t = \overline{r}$. The Foreign nominal interest rate is random. This setup implies that there are in effect only two assets: a risk-free Home bond and a risky Foreign bond whose return depends on nominal exchange rate fluctuations. I model the exogenous stochastic process for the interest rate differential $i_t - i_t^*$, defined as r_t, as following an unobserved components representation, whose details are described in section II.B.

Agents. — There are two types of agents in the economy. One type is the 'liquidity traders'. They have a passive role by holding every period a constant amount $\overline{X}$, expressed in USD, of Foreign bonds. The other type are 'investors'. There are overlapping generations of investors who each live two periods, derive utility from end-of-life wealth and are born with zero endowment. Investors born at time t are risk-neutral over end-of-life wealth, W_{t+1}, and face a convex cost of capital. Since PPP holds, Foreign and Home investors face the same real returns and choose the same portfolio.

Investment opportunity. — Investors pursue a zero-cost investment strategy: borrowing in one currency and lending in another. Investors make only one portfolio choice: the amount b_t, expressed in USD, of Foreign bonds invested. To illustrate the investment position suppose that b_t is positive. That means that the agent has borrowed b_t in USD and obtains $b_t \frac{1}{S_t}$ Foreign currency units (FCUs), where $S_t = e^{s_t}$. This amount is then invested in Foreign bonds and generates $b_t \frac{1}{S_t} \exp(i_t^*)$ of FCUs at time $t+1$. At time $t+1$, the agent has to repay the interest bearing amount of $b_t \exp(i_t)$. Thus, the agent has to exchange back the time $t+1$ proceeds from FCUs into USD and obtains $b_t \frac{S_{t+1}}{S_t} \exp(i_t^*)$. The net end-of-life realized wealth W_{t+1} is then, for a given choice b_t, a function of the realized excess payoff

$$W_{t+1} = b_t \exp(i_t)[\exp(s_{t+1} - s_t + i_t^* - i_t) - 1].$$

An approximation around the steady state of $i_t = i_t^* = 0$ and $s_{t+1} = s_t$ allows W_{t+1} to be written in the more tractable form as $W_{t+1} = b_t q_{t+1}$ where $q_{t+1} \equiv s_{t+1} - s_t - r_t$, denotes the log excess payoff.[7] The investor faces a convex cost of capital so that the realized payoff function for an investment b_t is $b_t q_{t+1} - \frac{c}{2} b_t^2$.

Market clearing condition. — There is a fixed supply B of Foreign bonds in the Foreign currency. In steady state, where variables are denoted by a subscript o, the investors demand zero Foreign bonds since there are no investment opportunities to take advantage of and they have zero endowment, i.e. $b_o = 0$. The real supply of Foreign bonds is $Be^{-p_t^*} = Be^{s_t}$ where the Home price level has been normalized to one. Thus, taking into account the demand from the liquidity traders, the market clearing condition is:

$$b_t + \overline{X} = Be^{s_t}. \tag{1}$$

Normalizing $s_o = 0$, it follows that $\overline{X} = B$. I set $B = 0.5$, corresponding to a two-country setup with half of the assets supplied by each country. For small s_t, the RHS of (1) becomes $B(1 + s_t)$. Under this approximation, I get the market clearing condition:

$$b_t = 0.5 s_t. \tag{2}$$

B. *Investor's optimization problem*

To model investor's preferences for uncertainty about the true DGP, I use the max-min expected utility axiomatized by Gilboa and Schmeidler (1989). These axioms imply that the ambiguity-averse investor behaves *as if* she maximizes

[7] I assume that investors have a large exogenous background income that is used to finance losses coming from the zero-cost strategy. Thus, there is no default risk on the part of the investor.

expected utility under a worst-case belief that is chosen from a *set* of conditional probabilities. This set is part of the primitives of the model. The *min* operator instructs the investor to identify the worst-case scenario that, conditional on the investment decision b_t, minimizes expected utility over elements of the set. The *max* operator means that the investor is maximizing this worst-case expected utility by choosing over the invested amount b_t of Foreign bonds. The investor's max-min expected utility at time t is:

$$V_t = \max_{b_t} \min_{\widetilde{P} \in \Gamma_t} E_t^{\widetilde{P}} [(b_t q_{t+1} - \frac{c}{2} b_t^2)], \tag{3}$$

where Γ_t is the set of possible probability distributions. In describing the set of beliefs Γ_t, I will discuss three aspects: first, in section II.B, I describe the true DGP and the type of ambiguity about it, namely uncertain time-varying signal to noise ratios. Second, in section II.B, I present the filtering problem implied by the uncertain precision of signals. Third, in section II.B, I discuss the conjectured law of motion for s_{t+1}.

Set of beliefs. — The agent observes the realized differential r_t, which is the sum of a hidden persistent component and a temporary shock. The stochastic process for r_t is:

$$r_t = x_t + \sigma_{V,t} v_t \tag{4a}$$

$$x_t = \rho x_{t-1} + \sigma_U u_t, \tag{4b}$$

where v_t and u_t are both Gaussian white noise. The true DGP is characterized by a sequence of constant variances that are equal to $\overline{\sigma}_V$ at each point in time:

$$\sigma_V^{DGP} = \{\sigma_{V,t} = \overline{\sigma}_V, \forall t\}. \tag{5}$$

As with a standard signal extraction problem, the agent does not observe separately x_t and v_t. The ambiguity aversion component is reflected in a further layer of incomplete knowledge: the agent is uncertain about the DGP and believes that $\sigma_{V,t}$ is potentially time-varying. The agent believes that the sequence of past realized variances $\{\sigma_{V,1}, ..., \sigma_{V,t}\}$ belongs to a set Γ_t of possible sequences. Here I conserve on notation and reuse Γ_t, which appeared in (3) and there denoted the set of possible probability distributions. Typical of ambiguity aversion frameworks, the agent's uncertainty manifests in her cautious approach of not placing a probability distribution over elements in this set. The agent entertains the possibility that any sequence out of the set Γ_t can be the true DGP. The agent trusts the remaining elements of the representation in (4).

I now describe the benchmark constraints that characterize the primitive set

Γ_t. This set is a constrained strict subset of $\Upsilon^t = \{\underbrace{\Upsilon \times ... \times \Upsilon}_{t \text{ times}}\}$, where Υ contains three elements: $\sigma_V^L \leq \overline{\sigma}_V \leq \sigma_V^H$, such that $\sigma_V^L \geq 0$. At each time t, the agent is only uncertain over *some* periods, indexed by a subscript $s \in \{1, ...t\}$, in which the realization $\sigma_{V,s}$ at that point in time s can be *any* draw out of Υ. The agent is then sure that the rest of the elements of the sequence $\{\sigma_{V,1}, ..., \sigma_{V,t}\}$ are equal to $\overline{\sigma}_V$. Specifically, for each history of dates up to and including time t, there are two restrictions on this sequence: (i) there is a constant number n of past dates for which the realization of σ_V at that date can be any draw out of Υ; (ii) these n dates can only occur in the last m observed periods, with $m \geq n$.[8]

Let a feasible alternative sequence $\{\sigma_{V,1}, ..., \sigma_{V,t}\}$ be denoted by $\widetilde{\sigma}_V^{(t)}$. To summarize the above two restrictions, the set Γ_t is comprised of sequences $\widetilde{\sigma}_V^{(t)}$ that have the generic form:

$$\widetilde{\sigma}_V^{(t)} = \{\underbrace{\overline{\sigma}_V, ..., \overline{\sigma}_V}_{t-m \text{ times}}, \underbrace{\sigma_{V,t-m+1}, ..., \sigma_{V,t}}_{\in P(m,n)}\} \tag{6}$$

where the set $P(m, n)$ denotes the set of possible sequences $\{\sigma_{V,t-m+1}, ..., \sigma_{V,t}\}$ such that: for n number of times any element of the set Υ can be chosen, and for the rest $m - n$ number of times only the element $\overline{\sigma}_V$ can be chosen. It follows that the number, denoted by y, of possible combinations of time indices that have length n is equal to[9]:

$$y = \frac{m!}{n!(m-n)!}. \tag{7}$$

Given that there are three elements in the set Υ, that means that for each sequence of length n of time indices, there are 3^n number of sequences of the form $\widetilde{\sigma}_V^{(t)}$. Since there are y possible combinations of time indices of length n, that means that the total number of sequences $\widetilde{\sigma}_V^{(t)}$ in the set Γ_t is equal to $3^n y$.

Using the sequence $\widetilde{\sigma}_V^{(t)}$ defined in (6), the optimization in (3) is:

$$V_t = \max_{b_t} \min_{\widetilde{\sigma}_V^{(t)} \in \Gamma_t} E_t^{\widetilde{P}}[(b_t q_{t+1} - \frac{c}{2} b_t^2)]. \tag{8}$$

I refer to the worst-case sequence $\widetilde{\sigma}_V^{(t)}$ that solves the minimization in (8) as the distorted sequence and to the probability distribution $\widetilde{P}$ over next period W_{t+1} implied by the distorted sequence as *distorted beliefs*.

[8]If the number of observations in the sample is smaller than m, i.e if $t < m$, then m is set to t.

[9]To see this, take for example $m = 3$ and $n = 2$. Then the possible time indices for which the agent can choose any value in Υ are: $(t-2, t-1)$, $(t-2, t)$ or $(t-1, t)$. Note also that if $m = n$, then there is only one possible sequence of time indices, namely $(t - n + 1, ..., t - 1, t)$.

I conclude this subsection by discussing the modeling choices of setting the parameters m and n. First, the restriction of setting $m \geq n$ is made purely for tractability. Setting m to a finite value caps the number of possible time indices at which the agent entertains values for σ_V out of the set Υ. For example if $m = t$, such that the n dates could have been any of the previous periods, then the number y in (7) is increasingly large.

Second, the restriction implied by n is motivated both by tractability reasons and by a desire to control how 'close' the set Γ_t is to the true DGP. The tractability consideration is implied by the fact that even when $m = n$, there are 3^n number of sequences in the set Γ_t. Setting n to a small number eases the computations over this set. At the same time, the parameter n also controls how statistically plausible are the beliefs about $\widetilde{\sigma}_V^{(t)}$ given the data that the agent observes. Evaluating this plausibility is important in gauging how far away are the distorted beliefs from the ones implied by knowing the true DGP.

To quantify the 'distance' of the set Γ_t to the true DGP, I use likelihood comparisons. Specifically, I compare the log-likelihood of an observed sample $\{r^t\}$ computed under the true, constant variance sequence, denoted by $L^{DGP}(r^t)$, and under the distorted sequence, denoted by $L^{Dist}(r^t)$. This comparison is used to evaluate how 'likely' does the agent perceive the observed sample $\{r^t\}$ to be under the two alternative views about the underlying DGP. On average, because the sample $\{r^t\}$ is generated under the true DGP, the difference between $L^{DGP}(r^t)$ and $L^{Dist}(r^t)$ is positive and increasing in the number n. For a given desired average statistical performance of the distorted model, i.e. an average difference between $L^{DGP}(r^t)$ and $L^{Dist}(r^t)$, and given the set Υ, this restricts the elements in the distorted sequence to be different from the true sequence for a constant number n of dates. Intuitively, if n is low then the two sequences will produce relatively close likelihoods even though the sample is increasing with t. This approach of setting n low can also be interpreted as the agent viewing the possible alternative realizations in $\{\sigma_{V,1}, ..., \sigma_{V,t}\}$ as "rare events" compared to the "normal" times in which $\{\sigma_{V,s} = \overline{\sigma}_V, \forall s \leq t\}$.[10]

Filtering. — Let $\sigma_V(r^t)$ denote the distorted sequence $\widetilde{\sigma}_V^{(t)}$ of the form (6) chosen at time t, to reflect the worst-case agent's belief about the sequence of variances $\{\sigma_{V,s}\}_{s \leq t}$ that has generated the data $\{r^t\}$ observed by the investor.[11]

[10] An alternative, related to the strategy employed in Epstein and Schneider (2007), would be to fix a significance level for the likelihood ratio test so that $L^{Dist}(r^t)$ is lower than $L^{DGP}(r^t)$ every period by some fixed amount, and allow the number of dates n to vary at each date t. I choose to work with a constant n mainly for tractability reasons but also to capture the idea that for some particular samples, the distorted model is statistically performing better.

[11] Note that if $b_t = 0$, the payoff in (8) is insensitive to evaluations about x_t. For the ambiguity agent's investment decision to matter in determining the equilibrium s_t, the position b_t should be able to adjust to shocks. This point highlights the importance of the market clearing condition in (2). There if, contrary to the maintained assumption, the Foreign bond supply would be fixed in USD and no other traders would be willing to take the other side of the transaction, then b_t will always be equal to zero in equilibrium.

The ambiguity-averse agent acts *as if* this sequence characterizes the true DGP. Thus, *after* this sequence has been chosen by the agent at date t, it is treated as a deterministic sequence in the standard Kalman filter. The latter delivers the best statistical estimate of the hidden state x_t.[12]

The notation for the elements of deterministic sequence used in filtering has to reflect that the worst-case distorted sequence $\sigma_V(r^t)$ is a function of date t information. To account for this I introduce the following notation: let $\sigma_{V,(t),s}$ denote the value for the standard deviation of the temporary shock that was believed at time t to have happened at time s. The subscript t in parentheses refers to the period in which the minimization takes place. The subscript $s \leq t$ refers to the period in the observed sample $0, ..., t$ at which the draw for the standard deviation is believed to have been equal to that particular value of $\sigma_{V,(t),s}$.[13]

To keep track of this notation I denote by $I_i^t = \{r_s, \rho, \sigma_U, \sigma_{V,(t),s}, s = 0, ..., i\}$ for $i \leq t$, the information set that the filtering problem has at time i by treating the sequence $\{\sigma_{V,(t),s}\}_{s=0,...,i}$ as known. Note that this sequence is optimally selected at date t. This notation highlights that the filtering problem is backward-looking. Let $\widehat{x}_{i,i}^t$ be the estimate of the time i hidden state, x_i, based on the sample $\{r_s\}_{s=0,...,i}$ and let K_i^t be the time i Kalman gain by treating the sequence $\{\sigma_{V,(t),s}\}_{s=0,...,i}$ as known. Then, as shown for example in Hamilton (1994), for $i \leq t$:

$$\widehat{x}_{i,i}^t = E(x_i|I_i^t) = \rho\widehat{x}_{i-1,i-1}^t + K_i^t(r_i - \rho\widehat{x}_{i-1,i-1}^t) \quad (9a)$$

$$\Sigma_{i,i}^t \equiv E[(x_i^t - \widehat{x}_{i,i}^t)^2|I_i^t] = (1 - K_i^t)(\rho^2\Sigma_{i-1,i-1}^t + \sigma_U^2) \quad (9b)$$

$$K_i^t = (\rho^2\Sigma_{i-1,i-1}^t + \sigma_U^2)[\rho^2\Sigma_{i-1,i-1}^t + \sigma_U^2 + \sigma_{V,(t),i}^2]^{-1}. \quad (9c)$$

Conjectured law of motion. — For the minimization in (8) the agent needs to understand what is the direction in which her expected utility depends on the sequences $\widetilde{\sigma}_V^{(t)}$. From the time t perspective the only unknown that the agent cares about is the expected value of s_{t+1}. To obtain the minimizing solution, the agent has a guess about the evolution of s_{t+1}. Let $f(\widehat{x}_{t+1,t+1}^{t+1}, r_{t+1})$ denote the conjecture about how the future s_{t+1} is formed. The particular conjecture is:

$$s_{t+1} = f(\widehat{x}_{t+1,t+1}^{t+1}, r_{t+1}) = a_1\widehat{x}_{t+1,t+1}^{t+1} + a_2 r_{t+1}, \quad (10)$$

where the coefficients a_1, a_2 are to be determined in equilibrium. This guess is the same as the one that will hold under a rational expectations version of the model,

[12]For a deterministic sequence $\sigma_V(r^t, b_t)$, the agent is interested in estimating the hidden state by using the filter which minimizes the mean square error of the estimate. Given the linear Gaussian setup, the optimal filter is the standard Kalman Filter.

[13]This notation allows for the belief about the realization of σ_V at date s, $s \leq t-1$ to be different at $t-1$ and t, i.e that $\sigma_{V,(t-1),s} \neq \sigma_{V,(t),s}$. This can be interpreted as an update, although not Bayesian in nature.

as shown later in section III.A. It suggests that the future exchange rate s_{t+1} will respond to the observed differential at $t+1$ and to the time $t+1$ estimate of the time $t+1$ hidden state.

The guess will have to be verified in equilibrium as being consistent with the actual law of motion of the economy. In other words, I am looking for a law of motion for s_t that is the same at each point in time. In order to address this consistency, I impose the following restriction on the agent's conditional probability model:

Assumption 1: $E_t^{\widetilde{P}}(\sigma_{V,(t+1),t+1}) = \overline{\sigma}_V$.

Thus, there are two assumptions about the way that the agent at time t is forming expectations about s_{t+1} : one is that she uses the linear function $f(.)$ in (10) to map time $t+1$ objects into s_{t+1}. The second, reflected in Assumption 1, is that $\overline{\sigma}_V$ equals the time t forecast of the object $\sigma_{V,(t+1),t+1}$, which is the time $t+1$ beliefs about the $t+1$ value of σ_V.[14]

C. Equilibrium concept

Let $\{r^t\}$ denote the history of observed interest differentials up to time t, $\{r_s\}_{s=0,...t}$. Denote by $\widehat{x}_{t,t}^t$ the Kalman filter estimate of the time t hidden state based on $\{r^t\}$ and formed using $\sigma_V(r^t)$. Let $f(.)$ denote a linear function of the form $f(\widehat{x}_{t+1,t+1}^{t+1}, r_{t+1}) = a_1\widehat{x}_{t+1,t+1}^{t+1} + a_2 r_{t+1}$ as in (10) that controls the conjecture about s_{t+1}, so that $s_{t+1} = f(\widehat{x}_{t+1,t+1}^{t+1}, r_{t+1})$.

DEFINITION 1: *An equilibrium will consist of a conjecture $f(\widehat{x}_{t+1,t+1}^{t+1}, r_{t+1})$, an exchange rate s_t^*, a bond demand b_t^* and a distorted sequence $\sigma_V(r^t)$ for $\{r^t\}$, $t = 0, 1, ...\infty$ such that agents at time t use the data $\{r^t\}$ and $\sigma_V(r^t)$ for the state-space (4) to form an estimate $\widehat{x}_{t,t}^t$ and a conditional probability distribution over r_{t+1} which satisfy:*

1. Optimality: given s_t^* and $f(\widehat{x}_{t+1,t+1}^{t+1}, r_{t+1})$, the bond demand b_t^* and the distorted sequence $\sigma_V(r^t)$ are the optimal solution for the max-min problem in (8).

2. Market clearing: given b_t^*, $\sigma_V(r^t)$ and $f(\widehat{x}_{t+1,t+1}^{t+1}, r_{t+1})$, the exchange rate s_t^* satisfies condition (2).

3. Consistency of beliefs: $s_t^* = f(\widehat{x}_{t,t}^t, r_t)$.

[14]One might think that the agent should also consider the worst case future path for the exchange rate. This, however, raises technical issues. If there is a constant ambiguity premium generated by the expected arrival of future ambiguous news, that premium should be reflected in (10). But since there is no discounting of the expected premium in the UIP equation $s_t = E_t^{\widetilde{P}} s_{t+1} - r_t$, then the solution implies that the premium is equal to zero. The model could potentially feature a time-varying ambiguity premium resulting from the future ambiguous news, but the solution method needs to be significantly more complicated to account for the fact that the future worst-case scenario evaluations fluctuate as the future investment positions change.

III. Model solution

A. Rational Expectations Model Solution

Before presenting the solution to the model, I first solve the rational expectations (RE) version. Let P denote the true probability distribution. By definition, in RE the subjective and the objective probability distributions coincide, i.e. $P = \widetilde{P}$. For ease of notation, I denote $E_t(X) \equiv E_t^P(X)$. Note that the RE assumption is a special case of the model, in which $n = 0$ in (6) such that Γ_t is singleton and the only entertained sequence is the true constant sequence as in (5). The RE version of the problem in (8) is:

$$V_t = \max_{b_t} E_t[(b_t q_{t+1} - \frac{c}{2} b_t^2)]. \tag{11}$$

Combining the market clearing condition (2) with the FOC of problem (11) I get:

$$s_t = \frac{E_t(s_{t+1} - r_t)}{1 + 0.5c}. \tag{12}$$

I call (12) the UIP condition in the RE version of the model. Setting $c = 0$ implies the usual risk-neutral version $s_t = E_t(s_{t+1} - r_t)$. It is also easy to see that the problem in (11) can accommodate risk aversion. If $c = \gamma Var_t q_{t+1}$, where γ is the coefficient of absolute risk aversion, and $Var_t q_{t+1}$ is a constant, I obtain the standard mean variance utility.

To solve the model, I take the usual approach of a guess and verify method. To form expectations agents use the Kalman Filter. Let $\widehat{x}_{i,j} \equiv E(x_i|I_j)$ and $\Sigma_{i,j} \equiv E[(x_i - E(x_i|I_j))^2]$ denote the estimate and the mean square error of the hidden state for time i given information at time j. Thus, the estimates are updated according to the recursion:

$$\widehat{x}_{t,t} = \rho \widehat{x}_{t-1,t-1} + K_t(r_t - \rho \widehat{x}_{t-1,t-1}) \tag{13a}$$

$$K_t = (\rho^2 \Sigma_{t-1,t-1} + \sigma_U^2)[\rho^2 \Sigma_{t-1,t-1} + \sigma_U^2 + \overline{\sigma}_V^2]^{-1} \tag{13b}$$

$$\Sigma_{t,t} = (1 - K_t)(\rho^2 \Sigma_{t-1,t-1} + \sigma_U^2), \tag{13c}$$

where K_t is the Kalman gain. Based on these estimates let the guess about s_{t+1} be the same linear function as in (10). Specifically:

$$s_{t+1} = a_1 \widehat{x}_{t+1,t+1} + a_2 r_{t+1}. \tag{14}$$

For simplicity, I assume convergence on the Kalman gain and the variance matrix $\Sigma_{t,t}$. Thus, I have $\Sigma_{t,t} \equiv \Sigma$ and $K_t \equiv K^{RE}$ for all t. Then, since $E_t r_{t+1} = \rho \widehat{x}_{t,t}$, combining the FOC in (12) with the guess in (14), leads to the solution:

$$a_2 = -\frac{1}{1+0.5c}, \; a_1 = -\frac{\rho}{(1+0.5c)(1+0.5c-\rho)}. \tag{15}$$

For the case of $c = 0$, the coefficients in (15) are $a_1 = -\frac{\rho}{1-\rho}$, $a_2 = -1$. These coefficients highlight the 'asset' view of the exchange rate: s_t is the negative of the present discounted sum of the interest rate differentials. If r_t is highly persistent, s_t reacts strongly to the estimate of the hidden state $\widehat{x}_{t,t}$ because this estimate is the best forecast for the highly predictable future interest rates. As introduced in (10), the same law of motion for s_t will be used in solving for the ambiguity aversion model, except that there the estimates of the hidden state will be formed under the potentially distorted beliefs.

B. Ambiguity Aversion Model Solution

The max-min problem in (8) means finding two objects: the worst-case belief $\sigma_V(r^t)$ and the optimal investment decision b_t^*. In section III.B I describe some properties of the equilibrium characterizing these two objects. I start by solving for a minimizing distorted sequence, denoted by $\sigma_V(r^t, b_t)$, that is conditional an investment position b_t. To do so, I use three monotonicity results: first, I present a conjecture about the sign of the direction in which the equilibrium s_{t+1} responds to time $t+1$ objects. Second, in Proposition 1, I show that the time t expected payoff is monotonic in the hidden state $\widehat{x}^t_{t,t}$, with the sign of the monotonicity being determined by the sign of b_t. Then, Corrolary 1 characterizes the monotonicity direction with respect to the time t element of the sequence $\sigma_V(r^t, b_t)$. While $\sigma_V(r^t, b_t)$ is conditional on a particular b_t, in Proposition 2 I then find the global solution of the optimal position b_t^*. Finally, section III.B uses these properties of the equilibrium and describes an algorithm for equilibrium determination.

Properties of the equilibrium. — Let $\sigma_V(r^t, b_t)$ denote the worst-case scenario belief, conditional on b_t, about the sequence of variances $\{\sigma_{V,s}\}_{s\leq t}$ that has generated the observed data. Through the Kalman filter, described in equations (9), $\sigma_V(r^t, b_t)$ can only affect the estimate $\widehat{x}^t_{t,t}$, which in turn controls $E_t^{\widetilde{P}}(r_{t+1})$. From the time t perspective the only relevant unknown is s_{t+1}. Thus, to solve for $\sigma_V(r^t, b_t)$, the agent has to form a guess about the monotonicity of s_{t+1} with respect to $E_t^{\widetilde{P}}(r_{t+1})$. That conjecture is defined in (10) as $s_{t+1} = a_1\widehat{x}^{t+1}_{t+1,t+1} + a_2 r_{t+1}$, where

$$\widehat{x}^{t+1}_{t+1,t+1} = (1 - K^{t+1}_{t+1})\rho\widehat{x}^{t+1}_{t,t} + K^{t+1}_{t+1} r_{t+1}$$

To solve for $\sigma_V(r^t, b_t)$, suppose the guess satisfies the following property:

CONJECTURE 1: *In equilibrium $a_1 < 0$ and $a_2 < 0$.*

Then

$$\frac{\partial s_{t+1}}{\partial r_{t+1}} = a_1 K_{t+1}^{t+1} + a_2.$$

Using that

$$E_t^{\widetilde{P}}(r_{t+1}) = \rho \widehat{x}_{t,t}^t = \rho \left[\widehat{x}_{t-1,t-1}^t + K_t^t (r_t - \rho \widehat{x}_{t-1,t-1}^t) \right], \quad (16)$$

together with $K_{t+1}^{t+1} \geq 0$ and Conjecture 1, it is then straightforward to establish the following:

PROPOSITION 1: *Given Conjecture 1, the expected excess payoff, $E_t^{\widetilde{P}} b_t(s_{t+1} - s_t - r_t)$, is monotonic in $\widehat{x}_{t,t}^t$. The monotonicity is given by the sign of $-b_t$.*

The property implied by Proposition 1 is that the worst-case evaluation of the hidden state depends on the sign of b_t, and not on its magnitude. The logic behind the proposition is simple. Suppose that the agent invests in domestic bonds by borrowing in foreign bonds (i.e. $b_t < 0$). The agent's expected excess payoff is then decreasing in the future depreciation of the home currency. Given Conjecture 1, a higher future depreciation of the home currency, i.e. a higher s_{t+1}, occurs in equilibrium when the future interest rate differential r_{t+1} is lower. As shown in (16), the expected r_{t+1} is increasing in the estimate of the hidden state $\widehat{x}_{t,t}^t$. Thus, when $b_t < 0$, the agent's expected excess payoff is increasing in $\widehat{x}_{t,t}^t$. Similarly, when $b_t > 0$, the payoff is decreasing in this estimate. Notice also that when $b_t = 0$, the expected payoff is equal to 0, and thus independent of the value of $\widehat{x}_{t,t}^t$.

The implication of Proposition 1 is that this minimization problem is to choose the sequence out of the feasible set Γ_t that produces the lowest estimate of the hidden state when $b_t < 0$ and the highest when $b_t > 0$. An analytic representation of the solution to $\sigma_V(r^t, b_t)$ can be stated with respect to its time t element, namely $\sigma_{V,(t),t}^2$, by taking as given the other elements in the sequence:

CORROLARY 1: *The expected excess payoff, $E_t^{\widetilde{P}} b_t(s_{t+1} - s_t - r_t)$, is monotonic in $\sigma_{V,(t),t}^2$. The monotonicity is given by the sign of $[b_t(r_t - \rho \widehat{x}_{t-1,t-1}^t)]$.*

Proof. See Appendix. The impact of $\sigma_{V,(t),t}^2$ on the expected excess payoff is given by the following intuitive mechanism. For illustration, suppose again that the agent invests in the domestic bonds. From Proposition 1, the agent is concerned that the estimate of the hidden state, $\widehat{x}_{t,t}^t$, is lower. In turn, the variance $\sigma_{V,(t),t}^2$ influences $\widehat{x}_{t,t}^t$ through its negative effect on the Kalman gain K_t^t (see (9)). The Kalman gain K_t^t affects the estimate $\widehat{x}_{t,t}^t$ through the incorporation of the current innovation (see (16)). Thus, to reflect a concern for a lower estimate $\widehat{x}_{t,t}^t$, the agent chooses to act *as if* the variance $\sigma_{V,(t),t}^2$ is larger when the innovation $(r_t - \rho \widehat{x}_{t-1,t-1}^t)$ is positive and as if $\sigma_{V,(t),t}^2$ is smaller if the innovation is negative.

The same intuition holds if $b_t > 0$. In that case the investor, concerned about a future foreign currency depreciation, chooses to incorporate the innovation $(r_t - \rho\widehat{x}^t_{t-1,t-1})$ in the opposite way. The decision rule for choosing $\sigma_{V,(t),t}$ implied by Corollary 1 is that

$$\sigma_{V,(t),t} = \sigma_V^H \text{ if } b_t(r_t - \rho\widehat{x}^t_{t-1,t-1}) < 0 \tag{17a}$$

$$\sigma_{V,(t),t} = \sigma_V^L \text{ if } b_t(r_t - \rho\widehat{x}^t_{t-1,t-1}) > 0. \tag{17b}$$

An interpretation of this decision rule is that investors react asymmetrically to news. If the agent decides to invest in the domestic currency, then increases (decreases) in the domestic differential are "good (bad) news" for this investment strategy. An ambiguity averse agent facing signals of ambiguous precision will then tend to underweigh good news by treating them as reflecting temporary shocks and overweigh the bad news by fearing that they reflect persistent shocks.

As evident from Proposition 1, the solution for the minimizing sequence presented so far is taking as given an arbitrary investment position b_t. To find the optimal investment decision b_t^*, first note that a property of the max-min optimization as in (8) is that it can have kinked or interior solutions. The possible sequences $\widetilde{\sigma}_V^{(t)}$ imply different probability distributions $\widetilde{P}$ to evaluate the excess payoff of investing in the foreign bond $q_{t+1} = (s_{t+1} - s_t - r_t)$. Let $\{\mu_j\}_{j=1,...J} \equiv E_t^{\widetilde{P}_j}(q_{t+1})$ denote the time t expected values of q_{t+1} under various probability distributions $\widetilde{P}_j$ implied by the feasible sequences $\widetilde{\sigma}_V^{(t)} \in \Gamma_t$. As discussed in section II.B the number of possibilities is $J = 3^n\frac{m!}{n!(m-n)!}$. To form these expected values, the investor uses each feasible sequence together with the conjectured law of motion for s_{t+1} in (10). The min operator can then be interpreted as the agent being concerned about which model indexed by j is the true model. Then the problem in (8) can be restated as:

$$V_t = \max_{b_t}\min_{\mu_j}(b_t\mu_j - \frac{c}{2}b_t^2). \tag{18}$$

PROPOSITION 2: *If there are values μ_j, μ_k such that $\mu_j\mu_k < 0$, then the global maximum in (18) is $b_t^* = 0$. Otherwise, the solution is $b_t^* = \frac{\mu_t^*}{c}$ where $\mu_t^* =$* $\arg\min_{\mu_j}\frac{\mu_j^2}{2c} = \min_{\mu_j}(|\mu_j|)$.

Proof: See Appendix. Intuitively, if all the implied $\{\mu_j\}_{j=1,...J}$ have the same sign, then the investor can take advantage of that situation by investing in the direction that makes the expected payoffs positive. In that case, the minimizing μ is the one with the lowest absolute value. However, if there are some μ_j for which the sign changes with j the agent faces a difficult situation: if she takes a positive (negative) position, i.e. $b_t > 0$ or $b_t < 0$, then there are sequences $\widetilde{\sigma}_V^{(t)}$ that imply a negative (positive) expected q_{t+1} leading to negative expected payoffs. In that situation the global solution to the maximization problem in (18) is $b_t^* = 0$.

Equilibrium determination. — Proposition 1 described how to find the minimizing sequence $\sigma_V(r^t, b_t)$. Then, Proposition 2 shows the optimal solution for b_t. The above intuition for b_t^* is reflected in the following steps of an algorithm for equilibrium determination.

1. First, guess that all the implied possible $\{\mu_j\}_{j=1,...J}$ are strictly positive. The guess does not specify the exact values of $\{\mu_j\}_{j=1,...J}$, but only their sign. Denote by μ_1 the minimum payoff in absolute value. Then, from Proposition 2, $sign(b_t) = sign(\mu_1)$.

2. Using the obtained sign of b_t, the sequence $\sigma_V(r^t, b_t)$ is computed using Proposition 1. Based on $\sigma_V(r^t, b_t)$, $\rho\widehat{x}_{t,t}^t = E_t^{\widetilde{P}}(r_{t+1})$. Using the guess for s_{t+1}, compute $E_t^{\widetilde{P}} s_{t+1}$ and

$$s_t = \frac{E_t^{\widetilde{P}} s_{t+1} - r_t}{1 + 0.5c}. \tag{19}$$

3. The "verify" step: the objects $b_t, \sigma_V(r^t, b_t), s_t$ are an equilibrium if $sign(s_t) = sign(b_t)$.[15] Then $\mu_1 = \mu_t^* = E_t^{\widetilde{P}} s_{t+1} - s_t - r_t$. If $sign(s_t) \neq sign(b_t)$, the conjectured starting point is not correct. This means that there are some sequences $\widetilde{\sigma}_V^{(t)}$ which imply a different sign for some of the $\{\mu_j\}_{j=1,...J}$ than the starting guess. In that case, suppose that at time t *all* the implied $\{\mu_j\}_{j=1,...J}$ are instead strictly negative and repeat steps 1 to 3.

4. If the new guess about $\{\mu_j\}_{j=1,...J}$ does not result in an equilibrium, then there exist at least two different sequences $\widetilde{\sigma}_V^{(t)}$ that imply μ_i, μ_j such that $\mu_i \mu_j < 0$. Then, from Proposition 2, the optimal decision is $b_t^* = 0$ and $s_t^* = 0$.[16] This concludes the algorithm.

The above algorithm is taking as given the guess in (10). The last element of the description of the equilibrium in section (II.C) is to check the consistency of this law of motion. In general, when $\overline{\sigma}_V > 0$, finding the coefficients a_1 and a_2 requires numerical solutions to minimize the distance between the conjectured law of motion and the actual law of motion that takes into account the distorted expectations and Assumption 1. In that general case, Conjecture 1 is verified by checking the signs of a_1 and a_2. Such a numerical procedure is detailed in the online appendix.

[15] Notice how this verifies the initial guess in step 1. From (19), the equilibrium $\mu_t^* = E_t^{\widetilde{P}} s_{t+1} - s_t - r_t$ has the same sign as b_t. Since μ_t^* is such that it minimizes $b_t \mu_j$ over $j \in \{1, ..., J\}$, it means that all the implied $\{\mu_j\}_{j=1,...J}$ have the same sign and that $\mu_1 = \mu_t^*$ is the minimum payoff in absolute value.

[16] Note that in this case the distorted sequence $\sigma_V(r^t)$ is not uniquely determined, since any of the possible sequences $\widetilde{\sigma}_V^{(t)}$ is an equilibrium.

A SPECIAL CASE OF NO TRUE TEMPORARY SHOCKS. — In this subsection, I present a specific case of the model, in which the true DGP is characterized by no temporary shocks, i.e $\overline{\sigma}_V = 0$. Note also that because $\overline{\sigma}_V \geq \sigma_V^L \geq 0$, this imposes the restriction that $\sigma_V^L = 0$. This special case will be a feature of the benchmark parameterization of the model. There are two main analytical results when $\overline{\sigma}_V = 0$.

LEMMA 1: *Suppose $\overline{\sigma}_V = 0$. Then, without loss of generality $m = n$ for the sequence defined in (6). This sequence simplifies to:*

$$\widetilde{\sigma}_V^{(t)} = \{\sigma_{V,s} \in \Upsilon \textit{ for } s = t-n+1, ...t \textit{ and } \sigma_{V,l} = \overline{\sigma}_V, \textit{ for } l < t-n+1\}.$$

Proof. See Appendix. The intuition behind Lemma 1 is the following. The estimate $\widehat{x}_{t,t}^t$ is a weighted average of all previously observed differentials $r_s, s \leq t$, with weights that are a function of the time-varying standard deviation $\sigma_{V,s} \in \Upsilon$. Now suppose that from the perspective of time t, time τ is the first period in the past in which $\sigma_{V,\tau} = 0$. By the Kalman gain formula $K_\tau = (\rho^2\Sigma_{\tau-1,\tau-1} + \sigma_U^2)[\rho^2\Sigma_{\tau-1,\tau-1} + \sigma_U^2 + \sigma_{V,\tau}]^{-1}$, the gain is $K_\tau = 1$. Intuitively, $K_\tau = 1$ means that the hidden state at time τ was exactly equal to the observed differential r_τ, i.e. that $x_\tau = r_\tau$. This perfect observation of the hidden state means that there is no more relevant information contained in the past observations $\{r_i\}_{i<\tau}$. Thus, these past observations receive zero weight in the estimate $\widehat{x}_{t,t}^t$. So, without loss of generality m can be set equal to n.

Lemma 1 implies that the sequences that the agent compares are the ones that have elements different from the constant sequence of zeros only in the last n periods. This means that the number of combinations of time indices that the agent looks at, given by the number y in equation (7), is equal to $\frac{m!}{m!(m-m)!} = 1$. Indeed, the only sequence of time indices that the agent looks at is $(t-n+1, ..., t-1, t)$. In this case, because there are only two possible values for $\sigma_V \in \Upsilon = \{0, \sigma_V^H\}$, there are only 2^n sequences of $\widetilde{\sigma}_V^{(t)}$ in Γ_t.[17]

The second result provides a closed form solution for the coefficients a_1, a_2. This simplifies significantly the computation of the equilibrium and allows for a direct comparison with the RE version of the model.

LEMMA 2: *Suppose $\overline{\sigma}_V = 0$. Then the equilibrium law of motion is:*

$$s_t = a_1\widehat{x}_{t,t}^t + a_2 r_t, \tag{20}$$

[17]In fact there are only $n+1$ distinct numbers for $\widehat{x}_{t,t}^t$ to compare. The reason is that once the value of σ_V corresponding to $t-s$, for $s = 0, ..., n-2$, in the sequence $(\sigma_{V,t-n+1}, ..., \sigma_{V,t-1}, \sigma_{V,t})$ is equal to zero, then the value of $\widehat{x}_{t,t}^t$ is the same for all values of σ_V with time indices smaller than $t-s$. Intuitively, a value of $\sigma_{V,t-s} = 0$ perfectly reveals the state at $t-s$, so any previous information is irrelevant. Since this irrelevance matters for $z = n-1$ nodes of the sequence, the number of distinct values of $\widehat{x}_{t,t}^t$ is $2^n\frac{1}{2^z} + z = n+1$.

where a_1, a_2 *are the same coefficients as in equations (15), which characterize the RE case.*

Proof. See Appendix.

C. The UIP regression

In this section I describe the UIP regression and how the ambiguity aversion model has the potential to explain the UIP puzzle. To relate to the UIP regression, I set $c = 0$, so that the agent is still ambiguity averse but otherwise risk neutral. First note that in the model UIP holds ex-ante under the one step ahead worst-case probability distribution $\widetilde{P}$:

$$E_t^{\widetilde{P}}(s_{t+1}) = s_t + r_t. \tag{21}$$

Define $\widetilde{\varepsilon}_{t+1}$ as an innovation to s_{t+1} that is zero mean under $\widetilde{P}$: $E_t^{\widetilde{P}}(\widetilde{\varepsilon}_{t+1}) = 0$. Then, the realized exchange rate $s_{t+1} = E_t^{\widetilde{P}}(s_{t+1}) + \widetilde{\varepsilon}_{t+1}$.

Let P denote the time t conditional distribution for the future s_{t+1} implied by: (i) the investor's equilibrium decision rule and (ii) the exogenous shocks being drawn from the true DGP. This P distribution controls how s_{t+1} behaves on average in this economy. The P distribution is different from $\widetilde{P}$ because in the latter the shocks are drawn from the distribution for r_{t+1} implied by the worst-case sequence $\sigma_V(r^t)$. The realized exchange rate can alternatively be expressed as:

$$s_{t+1} = E_t^P(s_{t+1}) + \varepsilon_{t+1}, \tag{22}$$

where ε_{t+1} is defined as an innovation such that $E_t^P(\varepsilon_{t+1}) = 0$. Let η_t denote the difference between the expected value of s_{t+1} under $\widetilde{P}$ and P : $\eta_t \equiv E_t^{\widetilde{P}}(s_{t+1}) - E_t^P(s_{t+1})$. Then substituting out $E_t^P(s_{t+1})$ in (22) by using η_t and the UIP condition in (21), I get

$$s_{t+1} - s_t = r_t - \eta_t + \varepsilon_{t+1}.$$

The typical UIP regression ignores the term η_t and consists of:

$$s_{t+1} - s_t = \beta r_t + \varepsilon_{t+1}. \tag{23}$$

The estimated UIP regression coefficient is then:

$$\widehat{\beta} = 1 - \frac{cov(\eta_t, r_t)}{var(r_t)}. \tag{24}$$

While there is no available closed form solution for $\widehat{\beta}$ in (24), the model is able to generate a negative UIP regression coefficient due to the positive covariance between r_t and η_t. The basic intuition is the following. The ambiguity averse investor is concerned about a high future depreciation of the investment currency. In equilibrium, investors then tend to underestimate, compared to the true DGP, the hidden state of the investment currency. They do so by taking a worst-case evaluation of the Kalman gain to be used in updating this state. The larger the observed investment differential is, the larger this underestimation is because the Kalman gain multiplies the perceived innovation. Thus, the higher the investment differential is, the larger is on average the difference between the expected depreciation of the investment currency under the equilibrium beliefs and under the true DGP. The ability of the ambiguity aversion model to potentially generate a negative $\widehat{\beta}$ can be contrasted with that of the rational expectations (RE) version. As shown in section III.A, the solution under RE with a mean-variance utility is $s_t = E_t(s_{t+1} - r_t)/(1 + 0.5c)$ with $c = \gamma Var(s_{t+1})$, where γ is the coefficient of absolute risk aversion. It then follows that

PROPOSITION 3: *In the RE case, the coefficient in the UIP regression (23) is:*

$$\widehat{\beta}^{RE} = \frac{1}{1 + 0.5c}\left[1 - \frac{0.5\rho cK}{(1 + 0.5c - \rho)}\right]. \tag{25}$$

Proof: See Appendix. Risk neutrality is obtained when $c = 0$, so $\widehat{\beta}^{RE} = 1$. When $c > 0$, $\widehat{\beta}^{RE} < 1$ due to a RE risk premium in that model. By setting $\overline{\sigma}_V = 0$, which implies that $K = 1$, I obtain the lower bound on $\widehat{\beta}^{RE}$, denoted by $\widehat{\beta}^{RE,L}$:

$$\widehat{\beta}^{RE} \geq \widehat{\beta}^{RE,L} = \frac{1 - \rho}{1 + 0.5c - \rho}. \tag{26}$$

The lower bound formula in (26) shows that $\widehat{\beta}^{RE}$, while lower than 1, cannot be negative in the RE version of the model. Its magnitude is investigated numerically in section IV.B.

IV. Results

A. Parameterization

The benchmark parameterization is summarized in Table 1. There are two main categories of parameters: one refers to the true DGP of the exogenous interest rate differential, given by the state space representation of (4) with the constant volatility sequence $\sigma_V^{DGP} = \{\sigma_{V,t} = \overline{\sigma}_V, \forall t\}$. The benchmark values for the true DGP, $\overline{\sigma}_V, \sigma_U$ and ρ are based on maximum likelihood estimates of such a state-space representation for a set of developed economies, as presented in Table B1. Consistent with these estimated values, in the benchmark parameterization

$\overline{\sigma}_V = 0$. Besides empirical plausibility, an advantage of setting $\overline{\sigma}_V = 0$ is that, as shown in Lemmas 1 and 2, in this case the equilibrium has simple analytical solutions. In section IV.B below, I present the model's implications for alternative values for $\overline{\sigma}_V$ and ρ.

Table 1—Benchmark specification

$\overline{\sigma}_V$	σ_V^H	σ_V^L	σ_U	ρ	$\overline{r}$	c	n	m
0	0.0025	0	0.0005	0.98	0	0	2	n

The second set of parameters refers to the feasible distorted beliefs. Given that $\overline{\sigma}_V = 0$, I set $m = n$, as justified by the generality result in Lemma 1, and also $\sigma_V^L = 0$. The benchmark parameters imply that the steady state Kalman weight on the innovation used to update the estimate of the current state is 1, 1 and 0.17 for the true DGP, the low variance and the high variance case respectively. While there is no independent evidence on σ_V^H and n, I evaluate their plausibility by computing the statistical distance between the true DGP and the equilibrium distorted model. Specifically, I compare the log-likelihood of a sample $\{r^T\}$ computed under the constant variance $(L^{DGP}(r^T))$ and under the equilibrium distorted sequence $(L^{Dist}(r^T))$. Table 2 reports some statistics for $L^{Dist}(r^T) - L^{DGP}(r^T)$, calculated for a sample of $T = 300$. Row (1) of Table 2 reports the results for $n = 2$, for which the average difference is around -1.4. Row (2) shows that when $n = 20$, the difference is around -15 which implies that in this case the distorted sequence results in an extremely unlikely interpretation of the data. For this reason, I restrict n to be a small number. Thus, the benchmark parameters generate equilibrium distorted beliefs that imply a similar interpretation of the observed data, when compared to the true, unknown to the agent, DGP.

Table 2—Likelihood comparison: $L^{Dist}(r^T) - L^{DGP}(r^T)$

	$Mean$	$St.dev.$	$\%positive$
(1)	-1.4	1.23	0.05
	(0.06)	(0.03)	(0.01)
(2)	-15.1	2.8	0
	(3.2)	–	–

Note: Likelihood difference for $T = 300$. Standard deviation of statistics is computed across $N = 1000$ simulations and shown in parantheses. The last column refers to the percent of simulations for which the difference is positive. The first two rows are for $n = 2$ and the last two for $n = 20$.

Finally, the real interest rate $\overline{r}$ is normalized to zero. The cost of capital c is set to zero to have UIP hold exactly under the RE version of the model. Then, according to Lemma 2, the resulting equilibrium coefficients a_1 and a_2 can be determined analytically and for the benchmark specification they are equal to $-\frac{\rho}{1-\rho} = -49$ and -1 respectively.

B. The UIP puzzle

Before presenting the results from the ambiguity aversion model, it is worth investigating as a benchmark the case of RE but with risk aversion. Equation (26) gives the lower bound $\widehat{\beta}^{RE,L}$ on the estimated $\widehat{\beta}$ in the UIP regression (23). To investigate the magnitude of this lower bound, I report below some simple calculations based on the main parameterization for the interest rate differential from Table 1. The equilibrium coefficients a_1 and a_2 are given by formulas (15). Table 3 reports the model implied $\widehat{\beta}^{RE,L}$ for various levels of risk aversion γ. With a low risk aversion, the model implied lower bound is smaller than one, but very close to it. Although $\widehat{\beta}^{RE,L}$ decreases with γ, even with a huge degree of risk aversion the coefficient is still positive and large. For example, when $\gamma = 500$ the value of the coefficient is around 0.41. Driving $\widehat{\beta}^{RE}$ to zero from above requires appealing to enormous levels of risk aversion.

Table 3—Rational expectations model

	$\gamma = 2$	$\gamma = 10$	$\gamma = 50$	$\gamma = 500$
$\widehat{\beta}^{RE,L}$	0.97	0.89	0.71	0.41

To present the numerical results for the ambiguity aversion model, I perform the following simulations. I first let the economy remain in a steady state in which the differential is always equal to zero. The economy is in steady state until there is convergence on the Kalman filter objects K and Σ obtained using at each point in time the value $\overline{\sigma}_V$. After convergence is obtained, I shock the economy for some number of periods T with random draws from the true DGP for the exogenous interest rate differential. To compute some statistics I then repeatedly draw such samples of size T.

Figure 1 is the histogram of the estimated $\widehat{\beta}$ in the UIP regression in (23) for the benchmark ambiguity aversion model for $N = 10000$ samples of $T = 300$ and $T = 3000$. It shows that the vast majority of the estimates are negative. The dispersion of these estimates is wider and negatively skewed for the small sample case of $T = 300$.

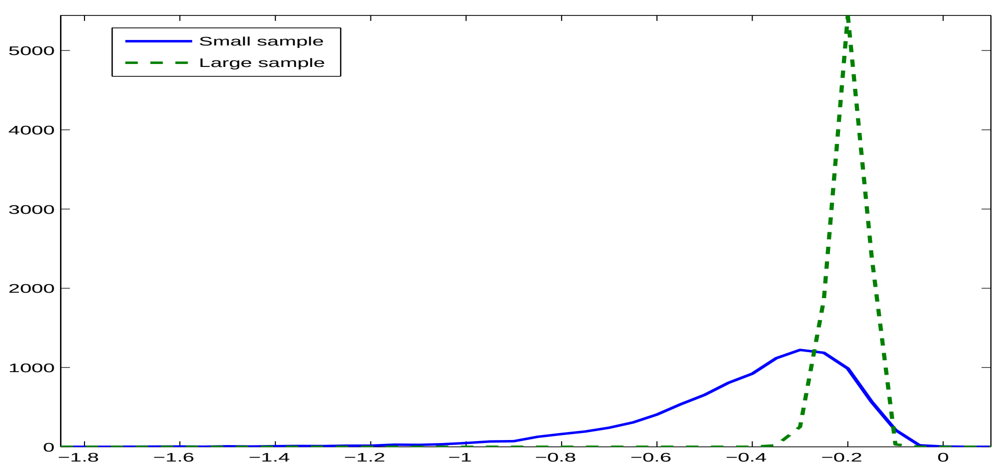
Table 4 reports some statistics for the estimated $\widehat{\beta}$ across various specifications. For each parameterization I report the value of $\widehat{\beta}$ and its t-statistic. Column (1) is the benchmark parameterization and shows that, for small T, the ambiguity aversion model is generating a negative average $\widehat{\beta}$, even if not significant statistically. As the sample size is increased and the standard errors reduce, the average estimate becomes significant. Reinforcing the message of Figure 1, these statistics highlight that the results of the model are not limited to small samples and in fact are stronger in large samples. The type of ambiguity modeled in this paper is active even as the sample size increases.

Table 4—Model implied UIP regression coefficients

		$\overline{\sigma}_V$=0	$\overline{\sigma}_V$=0.00055	$\overline{\sigma}_V$ =0.00125	ρ=0.7	$n = 20$
T=300	Mean	-0.42	-0.17	0.33	0.38	-0.6
		-0.81	-0.31	0.4	3.5	-1.3
	St.dev.	0.22	0.21	0.18	0.05	0.2
		0.38	0.25	0.24	0.4	0.5
T=3000	Mean	-0.2	-0.15	0.35	0.48	-0.3
		-1.7	-1.1	0.9	1.58	-2.2
	St.dev.	0.03	0.04	0.06	0.09	0.05
		0.17	0.19	0.18	0.2	0.2

Note: This table reports statistics for the estimated $\widehat{\beta}$ in the UIP regression in (23). Each entry has two values: the first refers to the value of the coefficient and the second to the value of the t-statistic of $\widehat{\beta}$.

Figure 1. Model implied UIP regression coefficients



Note: This figure is the histogram of the estimated $\widehat{\beta}$ in the UIP regression in (23) for the benchmark ambiguity aversion model for $N = 10000$ samples of $T = 300$ and $T = 3000$.

The benchmark parameterization is characterized by no temporary shocks and a large persistence of the hidden state. If the true DGP would feature much larger temporary shocks and much less persistence in the state evolution, the model would imply a positive, although still less than one, UIP coefficient. Interestingly, if one characterizes the developing economies as having these alternative properties for their interest rate process, then the model would be consistent with empirical observations, such as in Bansal and Dahlquist (2000), that the UIP coefficients tend to be positive for these countries.

The results for such alternative specifications are evaluated in Table 4. The main direction that is interesting to check is the quantitative role of $\overline{\sigma}_V$. Columns (2) and (3) consider cases in which the noise-to-signal ratio in the true DGP is increasingly higher than in the benchmark model. In Column (2), the only difference from the parameterization in Table 1 is that now $\overline{\sigma}_V = 0.00055$, which implies that the steady state Kalman gain for the true DGP, the high distorted precision and low distorted precision are $0.58, 1$ and 0.17 respectively. Thus, in this case, the distorted gains are relatively equally far from the gain implied by the true DGP. In this case, the UIP coefficients tend to be negative but closer to zero. In Column (3) $\overline{\sigma}_V = 0.00125$ and the steady state Kalman gain for the true DGP is equal to 0.3. Here the UIP coefficients are on average positive but not significantly different from zero.

Column (4) of Table 4 shows that when there is significantly less persistence in the state evolution the model cannot easily account for the UIP puzzle. For that experiment, the only change from the parameterization in Table 1 is that $\rho = 0.7$. The reaction of the exchange rate to the estimate of the hidden state is strongly affected by ρ because the present value of future payoffs to a bond is smaller following the same increase in the interest rate. For the same interest rate differential this makes agents demand less of the bond and the investment currency value goes up by less. A large sensitivity of the currency's value to the estimate of the hidden state allows small distortions to the estimate to produce large deviations in the exchange rate evolution. With significantly less persistence and reaction of the exchange rate, the model-implied UIP regression coefficient is smaller than one but positive.

Column (5) of Table 4 presents the results for the parameterization of the model in which there are much looser restrictions on the number of periods in which the agent entertains value different from $\overline{\sigma}_V$. Specifically, I set $m = n = 20$. For such a case, the model implies a negative and larger in magnitude UIP regression coefficient both in short and large samples. Thus, relaxing the benchmark restriction of having a small number n improves the model's ability to explain the empirical puzzles. However, as shown in Row (2) of Table 2, this improvement comes at the cost of having the distorted beliefs result in a very unlikely statistical interpretation of the data.

C. Ex-post carry trade payoffs

The UIP condition holds ex-ante under the equilibrium distorted beliefs so the equilibrium speculation strategy has zero expected payoffs under these beliefs. The ex-post failure of the UIP condition, described in Section IV.B, has significant consequences for the ex-post profitability of the investment strategy. As argued before, in equilibrium, the investors engage in the carry-trade strategy: borrow in the low interest rate currency and invest in the high interest rate currency. The payoff on a USD bet for the carry trade strategy is:

$$z_{t+1} = r_t - (s_{t+1} - s_t) \text{ if } b_t < 0 \tag{27a}$$

$$z_{t+1} = (s_{t+1} - s_t) - r_t \text{ if } b_t > 0. \tag{27b}$$

Using the definition of the carry trade in (27), Table 5 describes the model implied non-annualized monthly carry trade payoffs. These payoffs are characterized by a positive mean, negative skewness and excess kurtosis. The positive mean arises because a positive interest rate differential predicts on average a zero currency depreciation or even a slight appreciation of the high interest rate currency. The average mean payoff reported in Table 5 is 0.0018. To compare the model-implied

Table 5—Model implied statistics for the carry trade payoffs

	Mean	Standard Deviation	Sharpe Ratio	Skewness	Kurtosis
Data	0.0041	0.031	0.133	−0.26	4.25
$Model: T = 300$	0.0018	0.0171	0.1	−0.135	4.8
	(0.0008)	(0.001)	(0.048)	(0.25)	(0.7)
$Model: T = 3000$	0.0018	0.0172	0.1	−0.16	4.9
	(0.00002)	(0.00003)	(0.014)	(0.08)	(0.23)

Note: This table reports empirical and model-implied non-annualized monthly carry trade payoffs. Standard errors of the statistics across $N = 10000$ simulations are reported in parentheses.

statistics with the data, I compute the payoffs to the carry trade for 16 developed countries for the period 1976-2008. The main statistical properties of those payoffs are reported in Table B2 of Appendix B. There I report that the computed average mean payoff for the carry trade strategy is 0.0041.[18] Table 5 shows that the average standard deviation of the model implied payoffs is around 0.017. For the data analyzed in Table B2 the average standard deviation is 0.031. Thus, the model delivers mean payoffs that are around half of the empirical carry trade payoffs computed without transaction costs and at the lower bound of those with transaction costs. The model implied average standard deviation of the payoffs is around half of its empirical counterpart. With both the mean and the standard deviation lower, the model-implied carry trade payoffs have an average Sharpe ratio of 0.1, very close to the reported empirical value of 0.133.

[18]This does not take into account transactions costs. Burnside et al. (2011) analyze a more extensive data set and find that the average payoff to the carry trade without transactions costs across individual country pairs for the period 1976-2007 ranges from 0.0026 when the base currency is the GBP to 0.0042 when the base currency is the USD. With transaction costs they report a range of 0.0015 to 0.0025.

Besides the positive mean of the carry trade payoffs, a very interesting and important feature of the empirical payoffs is the negative skewness. Table 5 indicates that the model-implied payoffs to the carry trade are on average negatively skewed. The degree of skewness is slightly lower than in the data. The model implies a negative skewness of -0.135 while the average in Table B2 is -0.26.

To investigate further the properties of the realized skewness of excess payoffs, I construct two tests. The first, more cross-sectional in nature, is similar to that of Brunnermeier, Nagel and Pedersen (2008). It involves checking whether periods (countries in Brunnermeier, Nagel and Pedersen (2008)) characterized by a higher domestic currency also experience a negative skewness in the excess payoffs. To that end, I simulate the model for $T = 300$ and for each t I collect r_t and the subsequent realized excess payoffs $ex_{t+1} \equiv r_t - (s_{t+1} - s_t)$. I sort the excess payoffs ex_{t+1} according to the sign of r_t. Denote by ex^+_{t+1} the payoffs when $r_t > 0$ and by ex^-_{t+1} when $r_t < 0$. Consistent with the predictability of excess payoffs, the average of ex^+_{t+1} is positive and the average of ex^-_{t+1} is negative. Importantly, I find that the skewness of ex^+_{t+1} is negative and that of ex^-_{t+1} is positive.

The second test, for a time varying dimension, is to simulate the model and at each date t to take 20000 draws from the DGP process for the date $t + 1$ realizations of r_{t+1}. Using these draws, I solve the model at time $t + 1$ and then collect the equilibrium implied s_{t+1}. Based on these, I compute the realized excess payoffs ex_{t+1} and denote their skewness by $Skew_{t+1}$. I find that positive r_t are associated with negative $Skew_{t+1}$ and that a higher r_t predicts a lower $Skew_{t+1}$. In particular, in a regression of the form

$$Skew_{t+1} = \beta_2 r_t + \xi_{2,t+1},$$

I find an average $\widehat{\beta}_2$ of -2.2 with an average t-statistic of -6.8 for $T = 300$.

Consistent with the data, these results imply that investing in a high interest rate currency produces on average positive excess payoffs which are negatively skewed. The thicker left tail of the distribution of these payoffs occurs because of the larger reaction to negative innovations and the smaller reaction to positive shocks in the high-interest-rate. Thus "crash risk" in this model is endogenous and it happens when negative shocks hit an otherwise positive estimate of the hidden state.

D. Delayed overshooting

The UIP puzzle refers to the unconditional empirical failure of UIP. This does not necessarily imply that a conditional version of UIP fails too. Following a positive shock to the interest rate, the UIP condition states that the domestic currency should overshoot, by appreciating on impact, and then follow a depreciation path. Several studies have empirically investigated this impulse response using different identification restrictions and found evidence of delayed overshooting: following a contractionary monetary policy shock the domestic interest rate increases and there is a prolonged period of a domestic currency appreciation.[19]

To generate the impulse response of the exchange rate to a shock to the interest rate differential, I assume that the economy starts in steady state. Thus, $r_{t-1} = x_{t-1} = 0$, $b_{t-1} = 0$ and $E^{\widetilde{P}}_{t-1}s_t = E^{P}_{t-1}s_t = 0$. At time t, there is an observed increase of $\alpha > 0$ in r_t. Next periods shocks are all set equal to zero. To investigate the *average* response to an increase in r_t, this experiment needs to impose that the observed shock α is generated by a combination of a persistent and temporary shock that corresponds to their true DGP likelihood of occurrence. This means that, on average, given the observed increase of α, the *true* shock to the persistent hidden state x_t equals

$$x_t = E(\sigma_U u_t | \sigma_U u_t + \overline{\sigma}_V v_t = \alpha) = \alpha \frac{\sigma_U^2}{\overline{\sigma}_V^2 + \sigma_U^2}, \tag{28}$$

while the shock to temporary component is then $\alpha\overline{\sigma}_V^2/(\overline{\sigma}_V^2 + \sigma_U^2)$. The solution for the exchange rate is then: $s_t = a_1\widehat{x}_{t,t} + a_2 r_t$ where $\widehat{x}_{t,t}$ is the estimate of the hidden state x_t, formed either under rational expectations (RE) or the ambiguity aversion model.

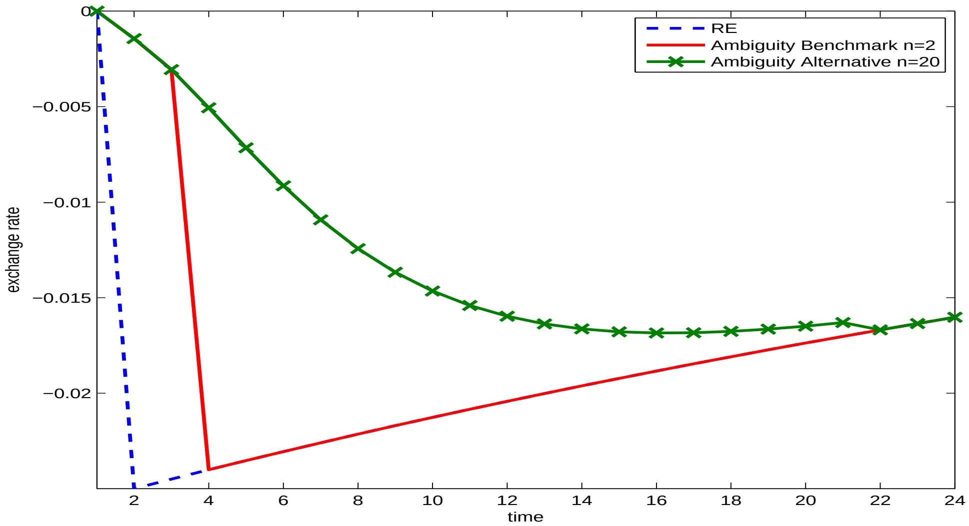
Figure 2 plots the average dynamic response of the exchange rate to the observed increase in r_t. There the shock occurs in period $t = 2$. The dashed (- -) line path is the RE model. This path features the Dornbusch (1976) overshooting so the peak of the response is at period 2. The solid (–) line path shows the evolution of s_t in the benchmark specification in which $n = 2$. There the peak response occurs 2 periods later than in the RE model. The starred (-x) line path considers the case of $n = 20$. The agent can distort any of the last 20 periods in the sample she observes, and in this case she does so by choosing a low precision of the signal for those 20 periods. In this case the appreciation is much more gradual and the peak is 15 periods later than the time of the shock.

The conclusion emerging from Figure 2 is that the model can qualitatively explain the delayed overshooting puzzle. The benchmark specification implies a quick peak and a short-lived deviation from UIP since the agent is limited by statistical plausibility considerations in distorting the time-varying precision of signals for only 2 periods. When these considerations are relaxed, the model delivers significantly longer delayed overshooting.

The intuition behind the delayed overshooting dynamic relies on the following key equilibrium observations: first, the agent undertakes the carry trade strategy by investing in the domestic bond. This means that she is worried about a significant future domestic currency depreciation. Second, given this investment

[19] Eichenbaum and Evans (1995) use short-run restrictions and find that the peak of the impact occurs after one to three years. Faust and Rogers (2003) find that these results are sensitive to the recursive identification assumptions and that the peak of the exchange rate response is imprecisely estimated. Scholl and Uhlig (2008) use sign restrictions and find that the estimated peak occurs within a year or two.

FIGURE 2. DELAYED OVERSHOOTING



Note: This figure plots the average dynamic response of the exchange rate to an observed increase in r_t of $\alpha = \sigma_U$. The shock occurs in period $t = 2$ and there are no further shocks. The dashed (- -) line path is the RE model. The solid (–) line path is the ambiguity benchmark specification of $n = 2$. The starred (-x) line is the case of $n = 20$.

position, the investor forms a cautious estimate of the hidden state controlling the domestic differential. She does so by overstating the probability that the underlying temporary shock has generated the observed increases in the domestic differential. The investor acts under the distorted belief that the realized variance equals σ_V^H and invests less in the domestic bond than she would under RE. Third, the systematic underestimation of the hidden state implies that the investor perceives on average positive innovations when updating the estimate. At that point, the positive perceived innovations create the possibility of further increases in the demand for the domestic currency leading to its gradual appreciation.[20]

E. Momentum-based investment strategies

The delayed overshooting experiment showed that following an increase in the investment differential, the investment currency experiences a gradual appreciation for several periods. Along that path, an econometrician would find evidence of time-series momentum: larger lagged ex-post excess payoffs predict larger future ex-post excess payoffs.

To investigate this type of strategy in my model, I will define the following

[20]This intuition can be presented more formally by investigating the analytical evolution of the estimate of the hidden state. Details of such an analysis are present in the online appendix.

modified carry trade which incorporates momentum:

$$z_{t+1} = r_t - (s_{t+1} - s_t) \text{ if } (r_t > 0 \text{ and } r_{t-1} - (s_t - s_{t-1}) \geq 0) \tag{29a}$$

$$z_{t+1} = (s_{t+1} - s_t) - r_t \text{ if } (r_t < 0 \text{ and } (s_t - s_{t-1}) - r_{t-1} \geq 0). \tag{29b}$$

The strategy described in (29) differs from the standard carry trade strategy, whose direction is defined only by the sign of r_t, by adding a momentum conditioning variable. This variable is the lagged excess payoff of the investment currency.

Table 6 reports empirical and model-implied results for the momentum-based strategy described in (29). In the data, this strategy produces a mean ex-post excess payoff that is about 70% larger than for the standard carry trade while generating similar volatility.[21] The model replicates qualitatively this empirical regularity. Rows (3) and (4) of Table 6 report the model implied payoffs. Compared to the standard carry trade, reported in Table 5, the momentum-based strategy generates payoffs and Sharpe ratios that are about 3 times larger. For this strategy, the model-implied payoffs are now much closer to the empirical ones, while for the standard carry trade they were about half as large.

Table 6—Statistics for the payoffs to the momentum-based strategy

	Mean	Standard Deviation	Sharpe Ratio	Skewness	Kurtosis
Data	0.0068	0.03	0.23	-0.29	4.1
	(0.0023)	(0.004)	(0.08)	(0.3)	(1.17)
Model	0.005	0.017	0.29	-0.09	4.58
	(0.0013)	(0.0012)	(0.075)	(0.3)	(0.77)

Note: This table reports empirical and model implied results for the momentum-based strategy described in (29). Rows (1) and (2) refer to the empirical average statistics across the 16 country pairs analyzed. Standard errors across these countries are in parentheses. Rows (3) and (4) report the model implied payoffs. Standard errors of the statistics across $N = 10000$ simulations are reported in parentheses.

F. Discussion on the structure of uncertainty

There are at least two features that need to be emphasized about the assumption that the agent is uncertain about the variance of the temporary shock $\sigma_{V,t}$, but otherwise trusts the remaining elements of the process in (4). First, is that, given ρ, what matters for the decision of the agent is the estimate of the hidden state.

[21] In Appendix B, Table B2 provides detailed statistics for each country pair. The last two columns refer to the momentum strategy. For 15 out of the 16 countries analyzed, the momentum-based strategy delivers higher payoffs than the standard carry trade. The profitability of time-series momentum strategies in the currency market has also been documented in Jorda and Taylor (2009), Moskowitz, Ooi and Pedersen (2011) and Burnside, Eichenbaum and Rebelo (2011). The latter find that these strategies are not easily explained by traditional risk factors.

This estimate is a function of the precision of the signal, which is controlled by the ratio of σ_U to σ_V. Thus, in analyzing uncertain precisions of signals, it is only this ratio that matters. For simplicity, I assume uncertainty only about σ_V, which can be interpreted as a normalization, but the same qualitative results go through if there is also uncertainty about σ_U.

Second, if there is uncertainty about ρ, a similar intuition would apply. To see that, take the following process with time-varying parameters:

$$r_t = \rho_t r_{t-1} + \sigma_{V,t} v_t \tag{30a}$$

$$\rho_t = \rho_{t-1} + \sigma_U u_t, \tag{30b}$$

where r_t is the observable interest rate differential, ρ_t is the hidden persistence parameter, v_t and u_t are both Gaussian white noise and $\sigma_{V,t}$ are draws from a set $\Upsilon_\rho = \{\sigma^L_{V,\rho}, \overline{\sigma}_{V,\rho}, \sigma^H_{V,\rho}\}$.

The solution for exchange rate determination for the process given by (30) is described in the online appendix. The basic intuition is the following: suppose that the agent observes a positive and larger than expected r_t. In equilibrium, the agent will want to take advantage of the positive r_t, but is concerned that the time t innovation is coming from a temporary shock and that the hidden ρ_t is low. The agent interprets news similarly to the benchmark model: a high (low) precision for negative (positive) innovations to the investment differential. Because ρ_t is underestimated, on average the observed r_{t+1} is higher than expected. That will create the possibility for a further appreciation of the high interest rate currency.

Finally, here I have assumed risk neutrality. Introducing risk-aversion complicates the analysis significantly. In that case, the minimizing variance of the temporary shocks has a second effect: a higher variance increases the expected variance of future payoffs. The overall effect of $\sigma^2_{V,(t),t}$ on the utility V_t is then coming through two channels.[22] From Corrolary 1, when there are "good news", the two effects align and the worst-case is a higher variance that will imply both lower expected payoffs and a larger variance of the payoffs. However, if there are "bad news", then the two directions are competing. Then it remains a quantitative question to determine which effect is stronger. To analyze this situation, I show in the online appendix that for a mean-variance utility, the benchmark specification implies that quantitatively the expected payoff effect is most likely to dominate.

[22]Illeditsch (2011) analyzes these two effects in a static environment.Li and Tornell (2008) show that if the agent is concerned only about uncertainty of the temporary (persistent) shock, then she will act as if the variance of that shock is higher, resulting in a robust Kalman gain that is lower (higher) than under RE.

V. Conclusions

This paper contributes to the theoretical literature that attempts to explain the observed deviations from UIP through systematic expectational errors. I present a model of exchange rate determination which features signal extraction by an ambiguity averse agent that is uncertain about the precision of the signals she receives. When deciding on the optimal investment position, the agent is estimating the time-varying hidden state of the exogenous observed interest rate differential. Faced with uncertainty, the agent chooses to act on beliefs that, compared to the true DGP, underestimate the hidden state of the investment differential, i.e. the differential between the higher interest rate and the lower interest rate.

Given the assumed structure of uncertainty, in equilibrium the agent systematically underestimates the hidden state by reacting asymmetrically to signals. The agent acts *as if* positive innovations to the investment differential, which in equilibrium are good news for the investor, are more likely to reflect a temporary shock, while negative innovations, which are bad news in equilibrium, as more likely to signal a persistent shock. The systematic underestimation implies that the agent perceives on average positive innovations when updating the estimate. This creates the possibility of a further increased demand next period for the investment currency and a gradual appreciation of it. Thus the model can provide an explanation for the UIP and the delayed overshooting puzzle.

The model provides a unified explanation for the main stylized facts of the excess currency payoffs: predictability, negative skewness and time-series momentum. Predictability is directly related to the ex-post failure of UIP. The negative skewness is caused by the asymmetric response to news. Time-series momentum is explained in the model by the gradual incorporation of good news in the estimate of the hidden state.

The theory proposed in this paper can be applied to other settings that involve forecastability of excess returns. Bacchetta, Mertens and van Wincoop (2009) use survey data to conclude that most of the predictability of excess returns in bond, stock and foreign exchange market is caused by predictability of expectational errors. Interestingly, in the stock market similar momentum strategies and impulse responses as the delayed overshooting puzzle have been documented (as for example in Hong and Stein (1999) and Moskowitz, Ooi and Pedersen (2011)). The model analyzed in this paper proposes dynamic filtering of signals with ambiguous precision as a mechanism to generate, through predictable systematic expectational errors, a positive mean, negative skewness and momentum effect for ex-post excess returns.

REFERENCES

Bacchetta, Philippe, and Eric van Wincoop. 2010. "Infrequent Portfolio Decisions: A Solution to the Forward Discount Puzzle." *American Economic*

Review, 100(3): 870–904.

Bacchetta, Philippe, Elmar Mertens, and Eric van Wincoop. 2009. "Predictability in Financial Markets: What Do Survey Expectations Tell Us?" *Journal of International Money and Finance*, 28(3): 406–426.

Bansal, Ravi, and Magnus Dahlquist. 2000. "The Forward Premium Puzzle: Different Tales from Developed and Emerging Economies." *Journal of International Economics*, 115–144.

Brunnermeier, Markus K., Stefan Nagel, and Lasse H. Pedersen. 2008. "Carry Trades and Currency Crashes." *NBER Macroeconomics Annual*, 23(1): 313–348.

Burnside, A. Craig, Martin S. Eichenbaum, Isaac Kleshchelski, and Sergio Rebelo. 2011. "Do Peso Problems Explain the Returns to the Carry Trade?" *Review of Financial Studies*, 24(3): 853–891.

Burnside, Craig, B. Han, D.A. Hirshleifer, and T.Y. Wang. 2010. "Investor Overconfidence and the Forward Discount Puzzle." *The Review of Economic Studies, forthcoming.*

Burnside, Craig, Martin S. Eichenbaum, and Sergio Rebelo. 2011. "Carry Trade and Momentum in Currency Markets." NBER Working Paper 16942.

Colacito, Riccardo, and Mariano M. Croce. 2011. "Risks for the Long Run and the Real Exchange Rate." *Journal of Political Economy*, 119: 153–182.

Dornbusch, Rudiger. 1976. "Expectations and Exchange Rate Dynamics." *Journal of Political Economy*, 84(6): 1161–76.

Eichenbaum, Martin, and Charles 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, Charles. 1996. "The Forward Discount Anomaly and the Risk Premium: A Survey of Recent Evidence." *Journal of Empirical Finance*, 3(2): 123–192.

Engel, Charles. 2011. "The Real Exchange Rate, Real Interest Rates, and the Risk Premium." NBER Working Paper 17116.

Epstein, Larry G., and Martin Schneider. 2007. "Learning Under Ambiguity." *Review of Economic Studies*, 74(4): 1275–1303.

Epstein, Larry G., and Martin Schneider. 2008. "Ambiguity, Information Quality, and Asset Pricing." *Journal of Finance*, 63(1): 197–228.

Fama, Eugene F. 1984. "Forward and Spot Exchange Rates." *Journal of Monetary Economics*, 14: 319–338.

Farhi, Emmanuel, and Xavier Gabaix. 2008. "Rare Disasters and Exchange Rates." NBER Working Paper 13805.

Faust, Jon, and John H. Rogers. 2003. "Monetary Policy's Role in Exchange Rate Behavior." *Journal of Monetary Economics*, 50(7): 1403–1424.

Frankel, Jeffrey, and Kenneth Froot. 1989. "Forward Discount Bias: Is It an Exchange Risk Premium?" *The Quarterly Journal of Economics*, 104(1): 139–61.

Froot, Kenneth A, and Richard H Thaler. 1990. "Anomalies: Foreign Exchange." *Journal of Economic Perspectives*, 4(3): 179–92.

Gilboa, Itzhak, and David Schmeidler. 1989. "Maxmin Expected Utility with Non-unique Prior." *Journal of Mathematical Economics*, 18(2): 141–153.

Gourinchas, Pierre-Olivier, and Aaron Tornell. 2004. "Exchange Rate Puzzles and Distorted Beliefs." *Journal of International Economics*, 64(2): 303–333.

Hamilton, James D. 1994. "Time Series Analysis." 372–408. Princeton University Press.

Hansen, Lars P. 2007. "Beliefs, Doubts and Learning: Valuing Macroeconomic Risk." *American Economic Review*, 97(2): 1–30.

Hansen, Lars P., and Robert J. Hodrick. 1980. "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis." *Journal of Political Economy*, 88(5): 829–53.

Hansen, Lars P., and Thomas J. Sargent. 2007. "Recursive Robust Estimation and Control without Commitment." *Journal of Economic Theory*, 136(1): 1–27.

Hansen, Lars P., and Thomas J. Sargent. 2010. "Fragile Beliefs and the Price of Model Uncertainty." *Quantitative Economics*, 1: 129–162.

Hong, Harrison, and Jeremy C. Stein. 1999. "A Unified Theory of Underreaction, Momentum Trading and Overreaction in Asset Markets." *Journal of Finance*, 54(6): 2143–2184.

Illeditsch, Philipp. 2011. "Ambiguous Information, Portfolio Inertia, and Excess Volatility." *Journal of Finance*, 66(6): 2213–2247.

Jorda, Oscar, and Alan M. Taylor. 2009. "The Carry Trade and Fundamentals: Nothing to Fear but FEER Itself." NBER Working Paper 15518.

Leippold, Markus, Fabio Trojani, and Paolo Vanini. 2008. "Learning and Asset Prices under Ambiguous Information." *Review of Financial Studies*, 21(6): 2565–2597.

Lewis, Karen K. 1989. "Changing Beliefs and Systematic Rational Forecast Errors with Evidence from Foreign Exchange." *American Economic Review*, 79(4): 621–636.

Li, Ming, and Aaron Tornell. 2008. "Exchange Rates Under Robustness: The Forward Premium Puzzle and Momentum."
Available at http://dx.doi.org/10.2139/ssrn.1300608.

Lustig, Hanno, and Adrien Verdelhan. 2007. "The Cross Section of Foreign Currency Risk Premia and Consumption Growth Risk." *American Economic Review*, 97(1): 89–117.

Lustig, Hanno, Nikolai N. Roussanov, and Adrien Verdelhan. 2011. "Common Risk Factors in Currency Markets." *Review of Financial Studies*, 24(11): 3731–3777.

Mark, Nelson C., and Yangru Wu. 1998. "Rethinking Deviations from Uncovered Interest Parity: the Role of Covariance Risk and Noise." *The Economic Journal*, 108(451): 1686–1706.

Moskowitz, Tobias, Yao H. Ooi, and Lasse H. Pedersen. 2011. "Time-Series Momentum." *Journal of Financial Economics*. forthcoming.

Plantin, Guillaume, and Hyun S. Shin. 2011. "Carry Trades, Monetary Policy and Speculative Dyamics." CEPR Discussion Paper DP8224.

Scholl, Almuth, and Harald Uhlig. 2008. "New Evidence on the Puzzles: Results from Agnostic Identification on Monetary Policy and Exchange Rates." *Journal of International Economics*, 76(1): 1–13.

Verdelhan, Adrien. 2010. "A Habit-Based Explanation of the Exchange Rate Risk Premium." *Journal of Finance*, 65(1): 123–145.

Appendix A: Proofs

Proof of *Corrolary 1*: From (9c):

$$\frac{\partial K_t^t}{\partial \sigma_{V,(t),t}^2} = -(\rho^2\Sigma_{t-1,t-1}^t + \sigma_U^2)[(\rho^2\Sigma_{t-1,t-1}^t + \sigma_U^2) + \sigma_{V,(t),t}^2]^{-2} < 0.$$

Thus

$$\frac{\partial \widehat{x}_{t,t}^t}{\partial \sigma_{V,(t),t}} = \frac{\partial \widehat{x}_{t,t}^t}{\partial K_t^t}\frac{\partial K_t^t}{\partial \sigma_{V,(t),t}^2} = \frac{\partial K_t^t}{\partial \sigma_{V,(t),t}^2}(r_t - \rho\widehat{x}_{t-1,t-1}^t).$$

Let $W_{t+1} \equiv b_t[s_{t+1} - s_t - r_t]$ denote the end-of-life wealth. Then using (16):

$$\frac{\partial E_t^{\widetilde{P}} W_{t+1}}{\partial \sigma^2_{V,(t),t}} = \frac{\partial E_t^{\widetilde{P}} W_{t+1}}{\partial \widehat{x}^t_{t,t}} \frac{\partial \widehat{x}^t_{t,t}}{\partial \sigma_{V,(t),t}}$$

$$sign\left[\frac{\partial E_t^{\widetilde{P}} W_{t+1}}{\partial \sigma^2_{V,(t),t}}\right] = sign\left[\frac{\partial E_t^{\widetilde{P}} W_{t+1}}{\partial \widehat{x}^t_{t,t}} \frac{\partial K^t_t}{\partial \sigma^2_{V,(t),t}} (r_t - \rho \widehat{x}^t_{t-1,t-1})\right].$$

Using Proposition 1 and the fact that $\partial K^t_t / \partial \sigma^2_{V,(t),t} < 0$, I get

$$sign\left[\frac{\partial E_t^{\widetilde{P}} W_{t+1}}{\partial \sigma^2_{V,(t),t}}\right] = sign\left[b_t(r_t - \rho \widehat{x}^t_{t-1,t-1})\right].$$

Proof of *Proposition 2*: Let $g_j(b_t) = (b_t\mu_j - \frac{c}{2}b_t^2)$ and $U(b_t) = \min_{\mu_j} g_j(b_t)$. Then (18) becomes $V_t = \max_{b_t} U(b_t)$. Compute the gradients of these functions at their only intersection point $b_t = 0$:

$$\frac{\partial g_j(b_t)}{\partial b_t}_{|b_t=0} = \mu_j.$$

If there are any μ_j, μ_k for which $\mu_j\mu_k < 0$, then the global solution is $b_t = 0$. If not, then the solution allows the interchangeability of the max and the min operator so that $V_t = \min_{\mu_j} \max_{b_t} g_j(b_t)$. For each model j we have the solution: $b_t = \frac{\mu_j}{c}$. So the minimization problem becomes $V_t = \min_{\mu_j} \frac{\mu_j^2}{2c}$ and the optimal solution is:

$$b_t^* = \frac{\mu_j^*}{c},\ \mu_j^* = \arg\min_{\mu_j} \frac{\mu_j^2}{2c} = \min_{\mu_j}(|\mu_j|).$$

Proof of *Lemma 1:* The recursion of the Kalman filter implies that:

$$\widehat{x}^t_{t,t} = (1 - K^t_t)\rho\widehat{x}^t_{t-1,t-1} + K^t_t r_t = K^t_t r_t + \sum_{j=1}^{t} \rho^j K^t_{t-j} r_{t-j} \left[\prod_{i=0}^{j-1}(1 - K^t_{t-i})\right].$$

Suppose that the first time going backwards in the sequence $t, ..., 0$ when the Kalman gain is equal to one is time τ, i.e. $K^t_\tau = 1$ and $K^t_s < 1$ for all $t \geq s > \tau$. Then:

$$\text{(A1)} \qquad \widehat{x}^t_{t,t} = K^t_t r_t + \sum_{j=1}^{t-\tau} \rho^j K^t_{t-j} r_{t-j} \left[\prod_{i=0}^{j-1}(1 - K^t_{t-i})\right].$$

Lemma 1 states that $\tau = t - n$. From (A1), any observation r_s with $0 < s < \tau$

receives zero weight in $\widehat{x}^t_{t,t}$, independent of $\sigma_{V,(t),s}$, $s < \tau$. Thus the minimization of $\widehat{x}^t_{t,t}$ over the n dates must involve choosing the dates n to be between τ and t.

Proof of *Lemma 2:* Take the conjectured law of motion in (10). Then:

$$E^{\widetilde{P}}_t s_{t+1} = E^{\widetilde{P}}_t[(1-K)a_1\rho\widehat{x}^{t+1}_{t,t} + (a_1K + a_2)r_{t+1}].$$

When $\overline{\sigma}_V = 0$, then $K = 1$. Use (19) and the same guess-and-verify method as for RE model to find that a_1, a_2 are given by (15).

Proof of *Proposition 3*: In the RE model the UIP regression (23) is

$$s_{t+1} - s_t = a_1(\rho - 1)\widehat{x}_{t,t} + a_1K(r_{t+1} - \rho\widehat{x}_{t,t}) + a_2(\rho\widehat{x}_{t,t} + \xi_t + \sigma_U u_{t+1} + \sigma_V v_{t+1} - r_t),$$

where $\xi_t = x_t - \widehat{x}_{t,t}$ with $\xi_t \sim N(0, \Sigma)$ and $E^P(\xi_t|I_t) = 0$. Since $E^P(\varepsilon_{t+1}|I_t) = 0$, then:

$$E_t(s_{t+1}) - s_t = -\widehat{x}_{t,t}\frac{0.5\rho c}{(1+0.5c)(1+0.5c-\rho)} + \frac{1}{1+0.5c}r_t.$$

Since $cov(\widehat{x}_{t,t}, r_t) = Kvar(r_t)$, the UIP coefficient in (23) is given by (25).

Appendix B: Supplementary tables

Table B1—ML estimates of state space representation

Belgium†	Canada	France†	Germany†	Italy†	Japan	Neth.†	Switz.	UK
0.97	0.96	0.95	0.99	0.98	0.99	0.99	0.99	0.98
(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
1.77	0.007	3.3	0.001	2.83	0.58	0.006	0.17	0.81
(0.7)	(1.08)	(1.02)	(0.41)	(1.02)	(0.38)	(0.45)	(1.58)	(0.53)
4.9	4.12	7.49	3.02	7.95	2.8	3.14	3.5	3.71
(0.5)	(0.17)	(0.83)	(0.16)	(0.74)	(0.2)	(0.17)	(0.25)	(0.28)

Note: The state-space representation is given by (4) with volatilities being constant. The values for ρ, σ_V and σ_U correspond to the first, third and fifth row. Standard errors for the estimated values are in parentheses. The entries for σ_V, σ_U are reported as their estimated values $\times 10000$. The sample is M1 1985-M12 2007 except for countries † for which data ends in M12:1998.

Source: Data is obtained from Datastream. The interest rate used is for a 1 month contract and it is the average of bid and ask quotes. The mnemonics for the data is, corresponding to the order of countries: Ecbfr1m, Eccad1m, Ecffr1m, Ecwgm1m, Ecitl1m, Ecjap1m, Ecnlg1m, Ecswf1m, Ecukp1m. To compute the interest rate differential with US, I use the corresponding Ecusd1m.

Table B2—Empirical UIP regression, carry trade and momentum payoffs

	UIP coeff.	Carry trade payoffs					Momentum	
Country	$\widehat{\beta}$	Mean	St.dev	SR	Skew	Kurtosis	Mean	St.dev
Austria†	-1.01	0.22	3.4	0.064	-0.14	0.78	0.47	3.2
	(0.73)	(0.22)	(0.2)	(0.067)	(0.17)	(0.39)	(0.25)	(0.2)
Belgium†	-0.66	0.69	3.3	0.208	-0.001	0.9	0.97	3.1
	(0.64)	(0.20)	(0.2)	(0.062)	(0.22)	(0.43)	(0.21)	(0.2)
Canada	-0.6	0.19	1.6	0.115	-0.5	1.27	0.19	1.6
	(0.5)	(0.09)	(0.1)	(0.055)	(0.16)	(0.46)	(0.11)	(0.1)
Denmark	-0.63	0.83	3	0.276	-0.13	0.84	1.1	2.8
	(0.47)	(0.17)	(0.1)	(0.059)	(0.14)	(0.41)	(0.17)	(0.17)
France†	0.06	0.53	3.2	0.168	-0.04	0.44	0.62	3.1
	(0.71)	(0.20)	(0.2)	(0.062)	(0.15)	(0.31)	(0.25)	(0.2)
Germany†	-0.66	0.12	3.4	0.035	-0.18	0.47	0.46	3.4
	(0.83)	(0.22)	(0.2)	(0.065)	(0.13)	(0.31)	(0.32)	(0.2)
Ireland†	0.38	0.53	3.2	0.165	-0.02	0.35	0.77	3.1
	(0.98)	(0.23)	(0.2)	(0.072)	(0.18)	(0.38)	(0.27)	(0.19)
Italy†	0.26	0.27	3	0.091	-0.33	1.11	0.74	3
	(0.4)	(0.21)	(0.2)	(0.07)	(0.24)	(0.57)	(0.2)	(0.3)
Japan‡	-2.55	0.28	3.5	0.08	-0.67	1.71	0.62	3.2
	(0.69)	(0.20)	(0.2)	(0.059)	(0.25)	(0.9)	(0.22)	(0.18)
Netherlands†	-1.68	0.35	3.4	0.103	-0.12	0.62	0.77	3.4
	(0.81)	(0.23)	(0.2)	(0.068)	(0.21)	(0.39)	(0.31)	(0.26)
Norway	-0.51	0.52	2.9	0.183	-0.12	1.12	0.67	3
	(0.5)	(0.14)	(0.1)	(0.05)	(0.17)	(0.43)	(0.2)	(0.18)
Portugal†	0.45	0.42	3.2	0.131	-0.07	2.38	0.71	3.2
	(0.25)	(0.21)	(0.2)	(0.065)	(0.38)	(0.97)	(0.27)	(0.27)
Spain†	0.75	0.32	3.2	0.102	-0.72	2.08	0.88	3
	(0.52)	(0.24)	(0.2)	(0.076)	(0.35)	(1.43)	(0.24)	(0.28)
Sweden	0.36	0.59	3	0.199	-0.78	3.23	0.88	2.8
	(0.69)	(0.15)	(0.2)	(0.058)	(0.35)	(1.48)	(0.15)	(0.28)
Switzerland	-1.40	0.09	3.5	0.025	-0.23	0.67	0.37	3.4
	(0.68)	(0.20)	(0.2)	(0.056)	(0.2)	(0.45)	(0.24)	(0.23)
UK	-1.67	0.57	3	0.191	-0.03	2.6	0.66	2.9
	(0.85)	(0.15)	(0.002)	(0.050)	(0.37)	(0.99)	(0.19)	(0.16)
Average	-0.57	0.41	3.1	0.133	-0.26	1.25	0.68	3

Note: The UIP regression is: $S_{t+1}/S_t - 1 = \beta(F_t/S_t - 1) + \varepsilon_{t+1}$. Standard errors are in parentheses. SR denotes the Sharpe Ratio. The mean and standard deviation of payoffs are in percentage points.
Source: The spot and forward rates are the average of bid and ask quotes. The Datastream mnemonics for the spot (forward) rate data is: Austsch (Austs1f), Belglux (Belxf1f), Cndollr (Cndol1f), Danishk (Danis1f), Frenfra (Frenf1f), Dmarker (Dmark1f), Ipunter (Ipunt1f), Italire (Italy1f), Japayen (Japyn1f), Guilder (Guild1f), Norkron (Norkn1f), Portesc (Ports1f), Spanpes (Spanp1f), Swekron (Swedk1f), Swissfr (Swisf1f). The data is then transformed using the pound/USD series Usdollr (Usdol1f). The sample is M1:1976 to M7:2008, except for †, which ends in M12:1998 and ‡ which begins in M7:1978.