

# Parameter Learning and the Carry Trade

Yichuan Wang\*

April 10, 2016

## Abstract

Volatile beliefs generate large and predictable carry trade returns. In a symmetric two country model with power utility, risk aversion of 5, and constant gain learning about mean consumption growth, the carry trade earns a Sharpe ratio of 0.21. I extend this basic model to accommodate correlated belief updates and Epstein-Zin preferences. While certain extensions manage to replicate both large carry trade and equity Sharpe ratios, they are unable to do so simultaneously with smooth exchange rates. My model has the empirical implication that if a group of countries have similar true mean growth rates, then a growth carry of sorting countries by recent consumption growth should earn excess returns on par with the carry trade. I find support for this in a panel of 14 developed world countries.

**JEL Classification:** D83, G12, G15

**Keywords:** Carry Trade, Subjective Beliefs, Exchange Rates, Parameter Learning

---

\*Wang: University of Michigan, [yichuanw@umich.edu](mailto:yichuanw@umich.edu). For helpful comments and discussions, I would like to thank Stefan Nagel and Martin Schmalz, as well as seminar participants at the Ross School of Business. All errors are my own.

# 1 Introduction

Parameter learning has been successful in explaining many closed economy asset pricing puzzles, such as the equity risk premium, return predictability, and excess volatility (Pastor and Veronesi, 2009; Lewellen and Shanken, 2002; Collin-Dufresne et al., 2015). Whereas agents in traditional rational expectations models know the true parameters governing the economy, agents in a learning model must use past data to infer the values of said parameters. Asset prices reflect agents' time varying and uncertain beliefs. Learning models are therefore able to match the reality of large, predictable, and highly volatile equity returns.

This paper explores how parameter learning can help resolve two open economy asset pricing puzzles: the carry trade and exchange rate smoothness.

1. The *carry trade puzzle* refers to the failure of uncovered interest rate parity (UIP) to hold in the cross section.<sup>1</sup> Investing in high interest rate currencies by borrowing in low interest rate currencies earns annual Sharpe ratios of around 0.50 after transaction costs (Lustig et al., 2011). Such a strategy is known as the carry trade. Moreover, carry trade returns are predictable using interest rate differentials. Backus et al. (2001) translate these facts into restrictions on stochastic discount factors (SDFs) in each country. They show that under conditionally log normal SDFs, rational expectations, and complete markets, carry trade conditional expected returns are equal to one-half the difference in the conditional variances of SDFs.<sup>2</sup> Therefore large and time varying carry trade returns imply highly volatile and heteroskedastic SDFs. But given that observed consumption growth is smooth and nearly homoskedastic, why are carry trade returns so large? Solutions to the equity premium puzzle that increase the volatility of SDFs are not enough to resolve the carry trade puzzle. Instead, one needs an explanation for large time varying differences in SDF volatilities across countries.
2. The *exchange rate smoothness puzzle* refers to the fact that exchange rates are smooth relative to the level of equity market Sharpe ratios. This fact implies that SDFs across countries must be highly correlated (Brandt et al., 2006). But given that consumption growth is nearly independent across countries, what is the common risk factor that causes SDFs to be correlated? Smooth exchange rates are also at odds with large carry trade returns. Intuitively, the SDFs that are volatile enough to match carry trade returns also cause exchange rates to be too volatile (Lustig and Verdelhan, 2016). Therefore it is important for models of the carry trade to also match smooth exchange rates.

---

<sup>1</sup>UIP fails in the time series as well. However, the asset pricing literature focuses on this cross sectional violation because it is easy to implement without estimating parameters about the time series relationship between forward discounts and currency depreciation.

<sup>2</sup>Under more general distributions for the SDFs, the carry trade premium becomes the difference in conditional entropies. The fundamental insight that the carry trade is about differences in higher order moments.

Rational expectations models resolve these puzzles by considering more sophisticated preferences and economies. [Verdelhan \(2010\)](#) uses a symmetric two country model with habit formation to generate the necessary heteroskedasticity in SDFs. However, his model fails to match exchange rate smoothness. [Bansal and Shaliastovich \(2012\)](#) use a two country model with heteroskedastic, correlated, long run risks to generate both heteroskedastic SDFs as well as smooth exchange rates. [Farhi and Gabaix \(2014\)](#) derive another resolution of the carry trade through disaster risks.

Beliefs open up a third dimension for explaining foreign exchange puzzles. Instead of relying on complicated preferences and consumption economies, I instead take a disciplined departure from rational expectations in the form of constant gain learning about mean consumption growth. There certainly exist rational expectations models that explain the carry trade and smooth exchange rates, but my goal is to show how time varying beliefs can offer a more parsimonious explanation of these puzzles.

I start with a simple two country model with power utility and independent consumption growth to outline the basic mechanisms for how beliefs can generate a carry trade. When agents have to learn about mean consumption growth they will end up perceiving exchange rate drifts that do not show up in the data, thereby leading to carry trade opportunities. Interest rate differentials then track the extent of these belief errors, and thus interest rate differentials end up predicting carry trade returns. Volatile beliefs also introduce a volatile common belief distortion term into the SDFs of different countries, thereby explaining the exchange rate smoothness puzzle.

Ideally, a model of beliefs and the carry trade should also be able to price equity risk. Such a model would demonstrate the essential unity of fluctuating beliefs as a way of explaining asset pricing anomalies. Moreover, the exchange rate smoothness puzzle started based on arguments about what equity Sharpe ratios implied about plausible levels of exchange rate smoothness. A model that can jointly match the equity premium, the carry trade, as well as the smoothness of exchange rates would be a model that more credibly addresses the exchange rate smoothness puzzle as well.

In order to generate large equity Sharpe ratios, I combine Epstein-Zin preferences with priced parameter uncertainty. A power utility model fails to generate large equity Sharpe ratios because the increase in parameter uncertainty caused by learning does not have a substantial effect on valuations. In contrast, Epstein-Zin preferences induce a preference for the early resolution of uncertainty. Uncertainty about mean consumption growth lowers equity valuations and raises wealth returns ([Collin-Dufresne et al., 2015](#)). The long run risks literature for international finance has made heavy use of *correlated* long run risks.<sup>3</sup> The analogous mechanism in a model of beliefs is to have correlated belief updates. Intuitively, one

---

<sup>3</sup>For example, [Colacito and Croce \(2011\)](#); [Bansal and Shaliastovich \(2012\)](#) use correlated long run risks in their models that generate both smooth exchange rates and large equity premia.

might imagine that the developed world shares a common fate when it comes to long run consumption growth, but that the exact fate itself is undetermined. I derive novel tools for modeling multivariate constant gain learning with arbitrary belief correlations and study the asset pricing implications. Although my simple model is unable to simultaneously replicate large carry trade Sharpe ratios, equity Sharpe ratios, and smooth exchange rates, understanding why sheds insight on the limitations of models of belief formation in explaining anomalies in international finance.

I then evaluate the empirical predictions of my beliefs based story of the carry trade. My model predicts that carry trade returns should be determined by differences in expected growth rates, where expected growth rates come from an exponentially weighted averages of past consumption growth. I find support for this hypothesis in a panel of 14 developed world countries. This stylized fact also distinguishes my story of the carry trade from the existing long run risks and disasters explanations.

## 2 A Simple Model of Parameter Learning and the Carry Trade

A simple two country Lucas model most clearly illustrates the effect of learning. Let lower case letters denote logs. Global consumption growth is iid and distributed

$$\Delta \mathbf{c}_{t+1} = \begin{pmatrix} \Delta c_{t+1} \\ \Delta c_{t+1}^* \end{pmatrix} \sim \text{MVN} \left( \begin{pmatrix} g \\ g \end{pmatrix}, \sigma^2 I \right) \quad (2.1)$$

Where stars denote foreign variables. Agents know the variance  $\sigma^2$  but not the true mean growth rates  $g$ .

Agents have power utility. This implies log SDFs of

$$m_{t+1} = \log \delta - \gamma \Delta c_{t+1} \quad (2.2)$$

Because agents do not know the true growth rate  $g$ , they instead form beliefs  $\tilde{g}_t$  about consumption growth in the domestic country through constant gain learning:

$$\tilde{g}_t = \alpha \Delta c_t + (1 - \alpha) \tilde{g}_{t-1} = \sum_{j=0}^{\infty} \alpha (1 - \alpha)^j \Delta c_{t-j} \quad (2.3)$$

with gain parameter  $\alpha$ . Symmetric equations hold for foreign preferences and beliefs.

To see where the name constant gain learning comes from, it is useful to compare equation 2.3 to the full Bayesian benchmark. In that case, the belief at stage  $t$  is just the sample mean, and so the updating

equation can be written as

$$\tilde{g}_t = \underbrace{\frac{1}{t}}_{\text{Gain}} \Delta c_t + \left(1 - \frac{1}{t}\right) \tilde{g}_{t-1} \quad (2.4)$$

In this case, the gain parameter at stage  $t$  is  $\frac{1}{t}$  and decreasing over time. As  $t \rightarrow \infty$ ,  $\tilde{g}_t \rightarrow g$  almost surely and learning stops.

In contrast, constant gain learning replaces the decreasing sequence  $\frac{1}{t}$  by a small constant  $\alpha$ . Therefore agents never stop learning about the true growth rate. There are three reasons to focus on this case:

1. It is a sensible rule for learning if the agent is concerned about structural shifts in the true growth rate  $g$ . In this case the agent will want to overweight recent data when forming expectations.
2. There is empirical support for this model of belief formation. In the macro literature, [Milani \(2007\)](#) finds that constant gain learning is able to dramatically improve the fit of DSGE models to match smooth inflation and consumption growth. [Malmendier and Nagel \(2015\)](#) use survey microdata and find that respondents seem to form expectations in a fully Bayesian fashion but only over the events they see in their own life time. Embedding this form of limited Bayesian learning in an overlapping generations framework produces a learning process, when averaged over generations, that closely approximates constant gain learning.
3. It is analytically useful. Beliefs evolve according to an  $AR(1)$  and therefore have a stationary, non-degenerate distribution for all time periods  $t$ . Thus the trading opportunities caused by learning do not fade away with time.

Given that consumption growth is independent, then the beliefs  $\tilde{g}_t, \tilde{g}_t^*$  are independent. Agents know that they are learning from this process. So even though agents know that the true volatility of consumption growth is  $\sigma^2$ , they have a slightly wider predictive density because of the uncertainty induced by the mean. In particular,

$$\begin{aligned} \text{Var}(\tilde{g}) &= \sum_{j=0}^{\infty} \alpha^2 (1 - \alpha)^{2j} \text{Var}(\Delta c_{t-j}) \\ &= \sum_{j=0}^{\infty} \alpha^2 (1 - \alpha)^{2j} \sigma^2 \\ &= \frac{\alpha^2}{1 - (1 - \alpha)^2} \sigma^2 \\ &= \frac{\alpha}{2 - \alpha} \sigma^2 \end{aligned} \quad (2.5)$$

Therefore under the subjective distribution domestic and foreign consumption growth  $\Delta c_{t+1}$  is distributed

$$\Delta c_{t+1} \sim \text{MVN} \left( \begin{pmatrix} \tilde{g}_t \\ \tilde{g}_t^* \end{pmatrix}, \tilde{\Sigma} \right)$$

$$\tilde{\Sigma} = \frac{\alpha \sigma^2}{2 - \alpha} \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} + \sigma^2 \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$$

Define  $\tilde{\sigma}^2 = (1 + \frac{\alpha}{2-\alpha}) \sigma^2$  to be the subjective variance of domestic consumption growth.

Agents price assets under the subjective Euler equation

$$1 = \tilde{\mathbb{E}} [M_{t+1} R_{t+1}] = \int \phi \left( \frac{\Delta c_{t+1} - \tilde{g}_t}{\tilde{\sigma}} \right) \times \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} R_{t+1} d(\Delta c_{t+1}) \quad (2.6)$$

where  $\phi(\cdot)$  is the density of a standard normal distribution. This pricing equation differs from the rational expectations benchmark because agents price assets based on their subjective expectation  $\tilde{g}_t$  for future consumption growth instead of the true growth rate  $g$ .<sup>4</sup> I use  $\tilde{\mathbb{E}}$  to denote the expectation under the agents' subjective beliefs. Based on this Euler equation, the risk free rate reflects the subjectively expected growth rate

$$r_{t+1}^f = -\log \delta + \gamma \tilde{g}_{t+1} - \frac{1}{2} \gamma^2 \tilde{\sigma}^2 \quad (2.7)$$

The solutions for the price-consumption ratio also reflect subjective expectations:

$$PC_{t+1} = \frac{k}{1 - k} \quad (2.8)$$

$$k = \delta \exp \left( (1 - \gamma) \tilde{g}_{t+1} + \frac{1}{2} (1 - \gamma)^2 \tilde{\sigma}^2 \right) \quad (2.9)$$

$$r_{t+1}^w = \log \frac{PC_{t+1} + 1}{PC_t} + \Delta c_{t+1} \quad (2.10)$$

I close by specifying foreign exchange markets. I assume complete markets.<sup>5</sup> Therefore exchange rates are determined by ratios of SDFs.<sup>6</sup> Let  $S_t$  denote the time  $t$  spot exchange rate, expressed in units of

<sup>4</sup>In this simple model, agents also ignore the effects of future parameter revisions on price. Ignoring future parameter revisions is a useful approximation in the case of power utility (Collin-Dufresne et al., 2015). Intuitively, because the power utility SDF depends only on one period's realized consumption growth, then parameter uncertainty does not make the SDF more volatile.

<sup>5</sup>One could also ask whether relaxing the complete markets assumption is a more productive way to resolve exchange rate puzzles. Lustig and Verdelhan (2016), however, show that incorporating market incompleteness to existing complete markets models is not sufficient to jointly produce smooth exchange rates, large carry trade premia, and exchange rates that are uncorrelated with consumption growth. Therefore it is still necessary to consider various complete market benchmarks and to then figure out which particular cases of market incompleteness can help bridge the remaining facts.

<sup>6</sup>For a short proof, note there are two candidate SDFs for pricing domestic returns:  $M_{t+1}$  and  $\frac{S_{t+1}}{S_t} M_{t+1}^*$ . The second SDF arises because  $\frac{S_{t+1}}{S_t}$  converts domestic returns into foreign returns, and since the foreign agent also satisfies her first order conditions

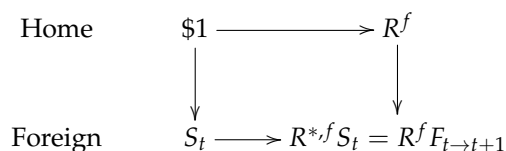
foreign currency per domestic currency. Then

$$\Delta s_{t+1} = m_{t+1} - m_{t+1}^* \quad (2.11)$$

Note that the foreign currency appreciates in bad times for the foreign investor and depreciates in good times for the foreign investor. Therefore when foreign consumption growth is high relative to domestic consumption growth, the foreign currency depreciates. Covered interest rate parity (CIP) governs the forward rate. Let  $F_{t \rightarrow t+1}$  denote the forward rate at time  $t$  for a contract that expires at time  $t + 1$ . That is, an agent, at time  $t$ , can lock in a contract to buy  $F_{t \rightarrow t+1}$  units of foreign currency per domestic currency in period  $t + 1$ . The law of one price implies that

$$f_{t \rightarrow t+1} = s_t + r_t^{f,*} - r_t^f \quad (2.12)$$

This condition comes about from analyzing two options that the agent has for investing one unit of domestic currency. In the first case, she can buy a forward contract, save the currency at home to have  $R^f$  in the home currency next period, and then convert into  $R^f F_{t \rightarrow t+1}$  units of foreign currency. Alternatively, she can convert today to get  $S_t$  units of foreign currency, and save that abroad at the interest rate of  $R^{f,*}$ . Since these two quantities can be locked in at period  $t$ , they must be equal, and so  $R^f F_{t \rightarrow t+1} = S_t R^{f,*}$ . Taking logs yields equation 2.12, and these relationships are illustrated in the diagram below.



My endowment model is robust to the introduction of production and trade costs. Assuming agents are optimizing their consumption, then the above equations are still first order conditions for the final consumption choices after agents have produced the optimal amounts of consumption and engaged in international trade. In fact, given that the exchange rate fluctuates, my model implicitly assumes that there are trade costs; otherwise purchasing power parity would hold and the exchange rate would be fixed at 1. But so long as agents optimize their consumption choices and financial markets are complete, consumption growth, interest rates, and exchange rates will be consistent with my model.

My model does assume that there is only one type of consumption produced in the world that enters

---

$\frac{S_{t+1}}{S_t} M_{t+1}^*$  becomes a candidate SDF for pricing domestic returns. Since SDFs are unique in complete markets, then  $M_{t+1} = \frac{S_{t+1}}{S_t} M_{t+1}^*$ . Equation 2.11 then follows from taking logs and rearranging terms. See Backus et al. (2001) for a more detailed explanation.

into both home and foreign utility functions additively. The exchange rate is the relative price of this consumption good at home versus consumption abroad.<sup>7</sup> There is no home bias in consumption, nor any special division between tradable and non-tradable goods. Rather, my model is a minimal one in which the real exchange rate is just the relative price of a consumption good that both domestic and foreign agents consume.<sup>8</sup>

## 2.1 Model Results

I simulate this model for 40,000 quarters. I calibrate consumption growth to the post 1980 developed world per capita consumption growth data, with an annualized mean of 1.51% and standard deviation of 1.67%. I set  $\delta = 0.995$  and  $\gamma = 5$ . I calibrate the gain parameter  $\alpha = 0.02$  to match the empirical estimates from Malmendier and Nagel (2015) and Milani (2007). Table 1 reports key moments from the simulation. Three themes emerge from the results:

### 2.1.1 Persistent Excess Returns Do Not Imply Compensation for Risk

One argument for risk based explanations of the carry trade is that the well known historical performance of the carry trade rules out stories of persistently biased agent expectations.<sup>9</sup> But constant gain learning means that even though beliefs about the growth rate of the economy are unconditionally unbiased, belief errors and carry trade profits never go away.<sup>10</sup> The returns section of table 1 reports large Sharpe ratios of 0.16 for a simple strategy of just buying the high interest rate currency, and a higher Sharpe ratio of 0.21 when the strategy scales with the size of the interest rate differential.<sup>11</sup> One way to conceptualize the magnitude of a 0.21 Sharpe ratio is to compare to Sharpe ratios typically seen in a power utility model with  $\gamma = 5$ . For example, my carry Sharpe ratio is more than twice the Sharpe ratio of the wealth return

---

<sup>7</sup>Equivalently, one unit of home consumption can be exchanged for  $S_t$  units of foreign consumption.

<sup>8</sup>Interpreting my model as story about two countries – the United States and the United Kingdom – can clarify my implicit assumptions about production and trade. Agents in both the US and UK have their own apple trees. Agents are indifferent between consuming American or British apples, and because it is costly to trade between the two countries, then the relative price of American and British apples fluctuates.

<sup>9</sup>Lustig and Verdelhan (2008) write “The average log excess return after transaction costs has increased from 227 basis points in the first part of the sample to 698 basis points per annum in the second part of the sample. Clearly, *this is not a temporary anomaly that is about to be arbitrated away*, and understanding carry trade returns is a critical step toward a better understanding of exchange rates.”

<sup>10</sup>My results that learning implies carry trade profits would also work in a fully Bayesian setting. So long as the beliefs are somewhat volatile, carry trade profits will exist. However, in that Bayesian setting profits would eventually go to zero as learning stops. Therefore constant gain learning allows me to more realistically model robust belief formation on the part of the agent as well as allowing profits opportunities to exist over time.

<sup>11</sup>This carry trade Sharpe ratios is lower than the 0.40 Sharpe ratio for going long/short *portfolios* of developed country currencies sorted by interest rates Lustig et al. (2011). They get larger Sharpe ratios because currency portfolios have lower volatility. But it is hard to compare their results with my result for two countries directly. One way to make a comparison is to suppose I had access to more currency pairs, thereby scaling expected returns by the number of additional countries and the exchange rate volatility at rate  $\sqrt{n}$ . Then if I go from 2 countries to 14 countries, as I do in my empirical analysis, then, assuming independent consumption, my Sharpe ratio would also be multiple by  $\frac{\sqrt{13}}{\sqrt{2}} \approx 2.54$ . My Sharpe ratio would then be more comparable. However, this is a very ad-hoc comparison, and I have a more thorough empirical test of my theory in section 4.



(0.08). As such it is important to understand why such large returns exist in equilibrium.

Analytically solving the carry trade premium reveals why the agents leave carry trade profits on the table. Buy the forward contract and sell it at the spot in the next period. The profits from this transaction are:

$$f_{t \rightarrow t+1} - s_{t+1} = -\Delta s_{t+1} + r_t^{f,*} - r_t^f \quad (2.13)$$

$$= m_{t+1}^* - m_{t+1} - \log \tilde{\mathbb{E}}_t [\exp(m_{t+1}^*)] + \log \tilde{\mathbb{E}}_t [\exp(m_{t+1})] \quad (2.14)$$

$$= m_{t+1}^* - m_{t+1} - \left( \tilde{\mathbb{E}}_t(m_{t+1}^*) + \frac{1}{2} \tilde{\text{Var}}_t(m_{t+1}^*) \right) + \left[ \tilde{\mathbb{E}}_t(m_{t+1}) + \frac{1}{2} \tilde{\text{Var}}_t(m_{t+1}) \right] \quad (2.15)$$

$$= [m_{t+1}^* - \tilde{\mathbb{E}}_t(m_{t+1}^*)] - [m_{t+1} - \tilde{\mathbb{E}}_t(m_{t+1})] \quad (2.16)$$

Equation 2.13 uses CIP. Equation 2.14 expands the exchange rate using equation 2.11 and the risk free rate using 2.7. Variance terms appear in equation 2.15 due to a Jensen inequality correction, and variances are eliminated from the last line because consumption variances, and thus SDF variances, are equal in the two countries.

To find the carry premium, I need to take expectations of equation 2.16. First, I take expectations under subjective beliefs. In that case

$$\tilde{\mathbb{E}} [f_{t \rightarrow t+1} - s_{t+1}] = \tilde{\mathbb{E}} \{ [m_{t+1}^* - \tilde{\mathbb{E}}_t(m_{t+1}^*)] - [m_{t+1} - \tilde{\mathbb{E}}_t(m_{t+1})] \} = 0 \quad (2.17)$$

This equation demonstrates that agents do not perceive any carry trade premium. On the other hand, when the econometrician measures the excess returns from the carry trade, she takes expectations under the objective distribution for consumption growth. In that case

$$\mathbb{E}_t [f_{t \rightarrow t+1} - s_{t+1}] = \mathbb{E}_t m_{t+1}^* - \tilde{\mathbb{E}}_t m_{t+1}^* - (\mathbb{E}_t m_{t+1} - \tilde{\mathbb{E}}_t m_{t+1}^*) \quad (2.18)$$

$$\gamma [(\tilde{g}_t^* - g) - (\tilde{g}_t - g)] \quad (2.19)$$

$$= \gamma (\tilde{g}_t^* - \tilde{g}_t) \quad (2.20)$$

$$= r_t^{f,*} - r_t^f \quad (2.21)$$

Thus the econometrician identifies a carry trade premium equal to the entire interest rate differential.

Agents fail to enter the carry trade because uncovered interest rate parity holds under the agents' subjective beliefs. High interest rate countries are those with high subjectively expected growth rates. But because of the positive relationship between a currency's strength and the value of the country's SDF, agents forecast currency depreciations for countries with high subjective growth rates. This forecasted

depreciation is precisely equal to the interest rate differential. On the other hand, the econometrician exploits the knowledge that consumption growth in the two countries share the same mean. Under the objective measure, SDFs are symmetric, exchange rates are a random walk, and the econometrician measures a carry trade premium that equals the interest rate differential. The carry trade exists not because agents are too risk averse, but rather because agents do not believe the carry trade will earn excess returns at all.

Another way to understand why agents are skeptical about carry trade profits comes from thinking about the carry trade returns that they must have observed as the two countries' interest rates  $r^f, r^{f,*}$  pulled apart. Suppose that  $r^{f,*} > r^f$ , so that there are large expected returns from investing in the foreign currency. By the Euler equation foreign growth expectations must also be higher than domestic growth rate expectations. For this to have occurred, it must have been that foreign growth consistently came in above the subjective expectation in the recent past. Given the tight link between growth and exchange rate depreciations imposed by equation 2.11, then it is also the case that foreign exchange rate depreciations came in larger than the interest rate differential in the recent past. During this time, the carry trade has consistently lost money. Constant gain learning then means this should affect the agents' future expectations, thereby making the agent unable to perceive the carry trade premium. Formally, we can decompose the current interest rate differential into past carry trade profits and past interest rate differentials:

$$\begin{aligned}
r_t^{f,*} - r_t^f &= \gamma (\tilde{g}_t^* - \tilde{g}_t) \\
&= \gamma [\alpha (\Delta c_t^* - \Delta c_t) + (1 - \alpha) (\tilde{g}_{t-1}^* - \tilde{g}_{t-1})] \\
&= -\alpha (r_t^{f,*} - r_t^f - \Delta s_t) + r_{t-1}^{f,*} - r_{t-1}^f \tag{2.22}
\end{aligned}$$

$$= -\alpha \sum_{i=0}^{j-1} (r_{t-i}^{f,*} - r_{t-i}^f - \Delta s_{t-i}) + r_{t-j}^{f,*} - r_{t-j}^f \tag{2.23}$$

Note that the term inside the parentheses is precisely the return from the carry trade. Therefore if the interest rate differential increased from period  $t - j$  to  $t$ , the term  $\sum_{i=0}^{j-1} (r_{t-i}^{f,*} - r_{t-i}^f - \Delta s_{t-i})$  must be negative, and so it must be that the carry trade lost money on average in periods  $t - j$  to  $t$ .

My explanation for the carry trade resembles [Lewellen and Shanken's](#) learning based explanation of the value premium (2002). In their model, agents learn about mean dividends in the cross section of stocks with identical mean dividend levels. Value stocks are stocks that have paid low dividends in the recent past, are perceived by agents to have low future dividends, and therefore have a low price. But because actual mean dividends have not decreased, expectations eventually mean revert, thereby making value

stocks appreciate in value. The value premium persists because agents cannot foresee this increase in dividend expectations in their subjective predictive distributions, and the book to market ratio tracks the difference between the agent's subjective belief and objective dynamics. Analogously, agents in my model would earn an excess return by buying currencies of countries with expected growth rates above the true growth rate. But the carry trade premium persists because the agents cannot know what the true growth rates are! Moreover, interest rate differentials track the magnitude of the difference between subjective beliefs and objective dynamics, and therefore forecast expected returns.

True, if agents were fully Bayesian and consumption truly had a constant mean, then learning would eventually stop and therefore no country would have a difference between the subjectively expected growth rate and the true growth rate. But my model has the general implication that so long as the agents' learning process is not fast enough to pin down the growth rate with absolute certainty, then carry trade premia can exist in the data without being compensation for risk.

## 2.1.2 Small Belief Distortions Can Look Like Large Risk Aversion

The carry trade Sharpe ratio is independent of risk aversion and linear in the difference in expected growth rates. Dividing equation 2.18 by the volatility of the exchange rate yields a carry trade Sharpe ratio of

$$SR_t [f_{t \rightarrow t+1} - s_{t+1}] = \frac{\tilde{g}_t^* - \tilde{g}_t}{\sigma\sqrt{2}} \quad (2.24)$$

Applying the law of iterated expectations also gives that the unconditional Sharpe ratio is equal to<sup>12</sup>

$$SR [f_{t \rightarrow t+1} - s_{t+1}] = \frac{\sqrt{2\alpha}}{\sqrt{\pi(2-\alpha)}} = \text{Std}(\tilde{g}_t) \times \frac{1}{\sigma} \sqrt{\frac{2}{\pi}} \quad (2.25)$$

Therefore large Sharpe ratios come from volatile beliefs caused by large gain and consumption volatility, not high risk aversion.

The closest rational expectations model that can generate a carry trade premium is an international Lucas model with heteroskedastic consumption growth. At time  $t$ , let next period consumption volatility at home and abroad equal  $\sigma_{t+1}^2, \sigma_{t+1}^{*2}$ , respectively. In that model the carry trade Sharpe ratio equals

$$SR_t [f_{t \rightarrow t+1} - s_{t+1}] = \frac{\gamma(\sigma_{t+1}^2 - \sigma_{t+1}^{*2})}{2\sqrt{\sigma_{t+1}^2 + \sigma_{t+1}^{*2}}} \quad (2.26)$$

<sup>12</sup>The conditional expected premium from going long the high interest rate currency is equal to  $|\gamma(\tilde{g}_t^* - \tilde{g}_t)|$ . Since  $\tilde{g}, \tilde{g}^*$  are independent and both normal with mean  $g$  and variance  $\frac{\alpha\sigma^2}{2-\alpha}$ , then  $\gamma(\tilde{g}_t^* - \tilde{g}_t) \sim N\left(0, \frac{2\gamma^2\alpha\sigma^2}{2-\alpha}\right)$ . By general results on half normals,  $E|\gamma(\tilde{g}_t^* - \tilde{g}_t)| = \frac{2\gamma\sigma\sqrt{\alpha}}{\sqrt{\pi(2-\alpha)}}$ . Then because the volatility of the strategy is always equal to  $\gamma\sigma\sqrt{2}$ , then the unconditional Sharpe ratio becomes the reported result.

A rational expectations econometrician facing a large unconditional carry trade premium, consumption variances that do not differ substantially across countries, and equation 2.26 must conclude that  $\gamma$  is very high.

On the other hand, under subjective beliefs I only need that the belief difference  $\tilde{g}_t^* - \tilde{g}_t$  to exhibit mild volatility. In fact, in my calibration,  $\tilde{g}_t$  is normally distributed with mean 1.51% and with standard deviation 0.16%. The full Bayesian solution for beliefs gives a good benchmark for conceptualizing how small belief fluctuations are in my model. The ratio between the volatility of beliefs and consumption volatility is  $\frac{\text{Std}(\tilde{g}_t)}{\sigma} = \sqrt{\frac{\alpha}{2-\alpha}} \approx 0.10$ . Hence for a fully Bayesian agent to have comparable belief volatility as the constant gain agent, the fully Bayesian agent would have to observe 25 years = 100 quarters of consumption data. This is a long time for consumption growth parameters to stay stable. Second, one can ask how quickly an agent with fully Bayesian belief formation would reject the constant gain belief. In particular consider the likelihood ratio  $L$  defined by

$$\begin{aligned} L &= \prod_{t=1}^T \frac{p(\Delta c_t | \tilde{g}_t)}{p(\Delta c_t | \bar{g}_t)} \\ &= \exp \left\{ -\frac{1}{2\sigma^2} \sum_{t=1}^T [(\Delta c_t - \tilde{g}_t)^2 - (\Delta c_t - \bar{g}_t)^2] \right\} \end{aligned}$$

Assume that the constant gain belief is initialized at  $\tilde{g}_0 = g$ , and that beliefs evolve according to equation 2.3. Figure 1 shows the value of  $\exp(\mathbb{E}L)$ , or the expected odds of the growth rate  $g$  equaling  $\tilde{g}_t$  instead of equaling the sample mean  $\bar{g}_t$ , as the number of observations  $t$  increases. This figure shows that the odds decline slowly. Even after 50 years of data, the agent can only say that the likelihood of the constant gain growth rate is 60% of the likelihood of the full Bayesian belief. This is not strong evidence for an agent concerned about robustness to reject the constant gain belief. Nor is there definite evidence preferring the sample mean over the constant gain learning belief even after 100 years, with odds only at 20%.

Subjective beliefs generate premia based on differences in expected first moments, whereas risk based explanations must depend on second moments and therefore the risk aversion parameter. In that sense my work is an extension of [Abel \(2002\)](#), who argued that one potential explanation of the equity risk premium was persistent pessimism on the part of agents. In that case, every percentage point of pessimism about consumption growth would raise the equity premium by an equal amount. Similarly, every percentage point difference in *relative* expectations about consumption growth in my model raises the conditional Sharpe ratio by an amount independent of  $\gamma$ . Ultimately, I can exploit mild volatility in beliefs in the place of large volatility in SDFs to generate large carry trade Sharpe ratios.

### 2.1.3 Subjective Beliefs Can Make Exchange Rates Smooth and Objective SDFs Highly Correlated

Allowing agents to have time varying subjective beliefs can dramatically change the interpretation of standard non-parametric tools used to evaluate rational expectations asset pricing models. One implication for international finance is that subjective beliefs can make measured SDFs based on [Hansen and Jagannathan \(1991\)](#) bounds look far more correlated than what seems to be justified on the basis of consumption growth. Equivalently, beliefs can explain how smooth exchange rates are consistent with large carry trade Sharpe ratios.

[Hansen and Jagannathan](#) bounds are non-parametric tools used to connect the levels of risk premia with the variance of marginal utility growth. Intuitively, for assets to earn large excess returns it must mean that agents are highly risk averse. Concretely, what this means is that the maximal attainable Sharpe ratio should be bounded above by the volatility of the agents' SDFs. Yet my basic model seems to violate this bound. The scaled carry trade earns a Sharpe ratio of 0.21 yet the volatility of the SDF is just  $\gamma\sigma \approx 0.08$ .

However, [Hansen and Jagannathan](#) bounds apply only to an SDF that prices returns under objective probabilities, whereas agents in my model price returns under subjective beliefs. Define an objective SDF  $\hat{M}_{t+1}$

$$\hat{M}_{t+1} = \frac{\tilde{f}_{t+1}}{f} \times M_{t+1} \quad (2.27)$$

Define the experienced marginal utility growth  $M_{t+1}$  as the subjective SDF. Rewrite the subjective Euler equation 2.6 under objective probabilities as:

$$1 = \tilde{\mathbb{E}} [M_{t+1}R_{t+1}] = \mathbb{E} \left[ \frac{\tilde{f}_{t+1}}{f} M_{t+1}R_{t+1} \right] = \mathbb{E} [\hat{M}_{t+1}R_{t+1}] \quad (2.28)$$

Where  $\tilde{f}_{t+1}/f$  is the ratio of the subjective belief and objective probability density functions. Equation 2.28 shows how the objective SDF earns its name: it prices returns under the objective probability distribution  $f$ . The objective SDF is the product of the belief distortion and the subjective SDF. [Cogley and Sargent \(2008\)](#), in their paper on the equity risk premium, emphasize that [Hansen and Jagannathan](#) bounds apply only to objective SDFs. Table 1 shows that while the [Hansen and Jagannathan](#) bound does not apply to the subjective SDF, which has a volatility of 0.08, it does apply to the objective SDF, which has a volatility of 0.30. Therefore high Sharpe ratios in a model with subjective beliefs indicate only that *either* beliefs or marginal utility growth must be volatile.

[Cogley and Sargent's](#) insight, when applied to international finance, can explain tightly correlated international SDFs and smooth exchange rates. Note that the tight connection between exchange rates

and SDFs in equation 2.11 implies

$$\underbrace{\text{Var}(\Delta s)}_{\text{Small}} = \underbrace{\text{Var}(m_{t+1})}_{\text{Large}} + \underbrace{\text{Var}(m_{t+1}^*)}_{\text{Large}} - 2 \underbrace{\text{Cov}(m_{t+1}, m_{t+1}^*)}_{\text{Large?}} \quad (2.29)$$

Given annual Sharpe ratios above 0.40 on international stock markets, Hansen and Jagannathan bounds imply that the stochastic discount factors within countries must have standard deviations of around 40%. Yet changes in log exchange rates have an annual standard deviation of only 10%. Therefore log stochastic discount factors must be highly correlated across countries.<sup>13</sup> But observable risk factors, such as consumption growth, have low correlations. These facts lead to the exchange rate smoothness puzzle: given plausible levels of correlation between SDFs, exchange rates in the data are too smooth (Brandt et al., 2006)

But because equation 2.29 applies to objective SDFs, it only describes moment restrictions on the combination of both the log subjective SDF and the belief distortion. Therefore the high covariation in SDFs implied by equation 2.29 can actually just be an artefact of a common time varying belief distortion. This belief distortion reflects the *common* deviation of beliefs shared by domestic and foreign agents from objective dynamics.

Studying the prices of Arrow securities on domestic and foreign outcomes shows how this common belief distortion has nothing to do with the relationship between shocks in the two countries. Let states be discrete, and consider the price of an Arrow security that pays out 1 unit of domestic consumption if a certain pair of domestic and foreign states  $(s, s^*)$  occurs. Domestic agents are willing to pay  $\tilde{f}_{t+1}(s, s^*) M_{t+1}(s)$ . Such a security pays out  $S_{t+1}$  units of foreign consumption, and therefore foreign agents would be willing to pay  $S_{t+1} \tilde{f}_{t+1}(s, s^*) M_{t+1}^*(s^*)$ . When the rational expectations econometrician computes what values the SDF takes in these states, she will divide by the objective probabilities and payoffs to get that the foreign SDF takes on a value of  $\frac{\tilde{f}_{t+1}}{f} M_{t+1}^* = \hat{M}_{t+1}^*$  and that the domestic SDF takes on a value of  $\frac{\tilde{f}_{t+1}}{f} M_{t+1} = \hat{M}_{t+1}$ . Both contain the common belief distortion  $\frac{\tilde{f}_{t+1}}{f}$ . This proof relied on complete markets and that domestic and foreign agents agree on the international distribution of consumption growth. Even if agents believe that foreign and domestic states are independent, the rational expectations econometrician will still recover SDFs that share a common belief distortion  $\frac{\tilde{f}_{t+1}}{f}$ .

Fortunately, the foreign belief distortion does not affect the price of all assets. In the case of independent shocks, the foreign belief distortion does not play a role in pricing assets based purely off of domestic shocks; Terran assets do not depend on belief distortions about Martian consumption growth. When pricing domestic payoffs under the objective measure, the foreign densities in the objective SDF decomposition

<sup>13</sup>Let  $m, m^*$  denote the domestic and foreign log SDFs. Then  $\text{Cov}(m, m^*) = \frac{1}{2} [\text{Var}(m) + \text{Var}(m^*) - \text{Var}(m - m^*)] = \frac{1}{2} (0.40^2 \times 2 - 0.10^2) = 0.155$ . Therefore  $\text{Corr}(m, m^*) = \frac{0.155}{(0.40)^2} = 0.97$ .

2.27 will factor out and have expectation 1 by construction. In the general case of correlated shocks, I can instead replace the ratio of foreign densities with the ratio of conditional distributions of foreign shocks, in which case they will again factor out and have no effect on assets based on purely domestic shocks.<sup>14</sup> For purely domestic assets, foreign belief distortions only matter to the extent that they convey information for domestic consumption shocks.

The common belief distortion term arises when one tries to find a pair of SDFs  $(M, M^*)$  that successfully prices international assets under the objective *joint* distribution represented by  $f_{t+1}$ . In this case the econometrician will instead pick up a common belief distortion about the joint distribution of shocks, thereby making objective SDFs highly correlated even when subjective SDFs are independent. This analysis is borne out in my simulations. The SDF moments section of table 1 shows that while the subjective SDFs are uncorrelated, the objective SDFs with the belief distortion term have a correlation of 0.92. Thus subjective beliefs help explain why exchange rates can be so smooth by changing what it means for objective SDFs to be correlated. Correlated objective SDFs need not reflect shared risk, but rather shared fluctuating beliefs.

### 3 Correlated Learning and Subjective Long Run Risks

I extend my model to include Epstein-Zin (EZ) preferences and correlated learning to find a way to match high equity Sharpe ratios as well. Agents in my model price a "subjective" long run risk: namely the risk that their beliefs about average consumption growth fluctuate over time. This extension to Epstein-Zin preferences helps to illustrate the limits of models of subjective beliefs in solving open economy asset pricing puzzles simultaneously with domestic asset pricing puzzles.

<sup>14</sup>Let  $s, s^*$  denote the shock variables for home and domestic. Let  $g(s)$  denote the marginal distribution of the home shock and  $g^*(s^*|s)$  be the density of the foreign distribution conditional on the domestic shock. Therefore the joint distribution is  $g(s) \cdot g^*(s^*|s)$ . Let tildes denote subjective belief distributions. Then if the payoff  $X_{t+1}$  depends only on  $s$ , then

$$\begin{aligned}
\mathbb{E} \left[ \frac{\tilde{g}(s)}{g(s)} \cdot \frac{\tilde{g}^*(s^*|s)}{g^*(s^*|s)} \cdot M_{t+1} X_{t+1} \right] &= \int \frac{\tilde{g}(s)}{g(s)} \cdot \frac{\tilde{g}^*(s^*|s)}{g^*(s^*|s)} \cdot M_{t+1} X_{t+1} P(ds \times ds^*) \\
&= \int \int \frac{\tilde{g}(s)}{g(s)} \cdot \frac{\tilde{g}^*(s^*|s)}{g^*(s^*|s)} \cdot M_{t+1} X_{t+1} g^*(s^*|s) g(s) ds^* ds \\
&= \int \tilde{g}(s) M_{t+1} X_{t+1} \left( \int \tilde{g}^*(s^*|s) ds^* \right) ds \\
&= \int \tilde{g}(s) M_{t+1} X_{t+1} ds = \mathbb{E} \left[ \frac{\tilde{g}}{g} \tilde{M}_{t+1} X_{t+1} \right]
\end{aligned}$$

### 3.1 A Model of Correlated Learning

I first derive formulas for correlated multivariate constant gain learning. My strategy is to show how the one dimensional version of constant gain learning can be derived from Bayesian preliminaries, and to then adapt this strategy for the multivariate case.

Let the agent have a normal prior for the mean growth rate  $g$  with a mean of  $\mu$  and variance  $\tau^{-2}$ . She samples consumption growth  $\Delta c_t$  from a normal distribution with unknown mean but known variance  $\sigma^2$ . The agent then forms a sequence of posterior means  $\{\tilde{g}_t\}$  for the mean growth rate that can be defined recursively:

$$\tilde{g}_t = \frac{1}{t + \sigma^2\tau^2}\Delta c_t + \left(1 - \frac{1}{t + \sigma^2\tau^2}\right)\tilde{g}_{t-1} \quad (3.1)$$

$$\tilde{g}_0 = \mu \quad (3.2)$$

To arrive at the constant gain learning formula 2.3, I first change the sample size  $t$  to the effective sample size  $\alpha^{-1}$  implied by constant gain learning, and then let the precision  $\tau \rightarrow 0$ . This establishes a model of constant gain learning that is independent of the prior.

I use the same strategy in the multivariate case. Impose a prior on the vector of mean growth rates  $\mathbf{g}$  with a mean of  $\boldsymbol{\mu}$  but with prior variance  $\mathbf{Y} = \tau^{-2} \begin{pmatrix} 1 & \phi \\ \phi & 1 \end{pmatrix}$ . In the multivariate case, there is a new parameter to specify – the correlation  $\phi$  of beliefs in the prior. A value close to 1 communicates the idea that the agent strongly believes the true means for consumption growth in the two countries are close together and therefore induces correlated belief updates. The agent samples consumption growth  $\Delta \mathbf{c}_t$  drawn from the multivariate normal distribution with unknown mean but known variance  $\Sigma = \sigma^2 I$ . The full Bayesian solution implies that the posterior means  $\{\tilde{\mathbf{g}}_t\}$  for the mean growth rate evolve as follows:

$$\tilde{\mathbf{g}}_t = M\Delta \mathbf{c}_t + (I - M)\tilde{\mathbf{g}}_{t-1} \quad (3.3)$$

$$M = \left(\Sigma \mathbf{Y}^{-1} + tI\right)^{-1} \quad (3.4)$$

$$\tilde{\mathbf{g}}_0 = \boldsymbol{\mu} \quad (3.5)$$

Replace  $t$  with  $\alpha^{-1}$  and let the precision go to zero. Appendix C shows that for fixed values of  $\phi < 1$ , the updating equation reduces to independent learning.<sup>15</sup> As precision  $\tau \rightarrow 0$ , the agent puts no weight on

<sup>15</sup>For  $\phi = 1$ , the prior corresponds to a dogmatic belief that mean growth rates in the two countries are equal, and is thereby successful at replicating perfectly correlated beliefs. However, since I want to be able to model beliefs with arbitrary correlation in  $(0,1)$ , I develop further methods.



the prior. Therefore  $\phi$  does not matter as the known independence in realized consumption overwhelms any motivation to update beliefs in a correlated fashion.

I resolve this problem by instead modeling the belief correlation  $\phi$  with local to unity asymptotics  $\phi \sim 1 - c\tau^2 + o(\tau^2)$  for some  $c > 0$ , and then letting  $\tau \rightarrow 0$ . Intuitively, this allows me to impose a strong prior that growth rates are correlated without also determining what each of the growth rates should be. The particular value of the hyperparameter  $c$  can be chosen to match any arbitrary belief correlation between 0 and 1. Appendix C outlines the derivations of the updating formulas and the theoretical variances and correlations of beliefs. The resulting update equations take the form

$$\tilde{\mathbf{g}}_t = M\Delta\mathbf{c}_n + (I - M)\tilde{\mathbf{g}}_{t-1} \quad (3.6)$$

$$M = \left( \frac{\sigma^2}{2c} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \alpha^{-1}I \right)^{-1} \quad (3.7)$$

$$\tilde{\mathbf{g}}_0 = \boldsymbol{\mu} \quad (3.8)$$

Although equation 3.6 indicates that the domestic agent will use realizations of foreign consumption growth to update beliefs, she nonetheless perceives that her subjective beliefs will evolve as a martingale. An important corollary is that future beliefs about domestic consumption,  $\tilde{g}_{t+1}$ , are independent of current beliefs about foreign consumption,  $\tilde{g}_t^*$ . Therefore the pricing of the domestic consumption claim does not depend on beliefs about foreign consumption growth.

To show that  $\tilde{g}_{t+1} \perp \tilde{g}_t^*$ , observe that the future belief is a normal random variable. The variance is already pinned down by known parameters, so to show that the two variables are independent, it suffices to show that  $\mathbb{E}_t[\tilde{g}_{t+1}] \perp \tilde{g}_t^*$ . There are two ways to arrive at this result. First, the evolution of future beliefs under current beliefs must be a martingale. Otherwise the agent would have revised her current period belief. Therefore  $\mathbb{E}_t[\tilde{g}_{t+1}] = \tilde{g}_t$ , which is already determined and therefore does not depend on  $\tilde{g}_t^*$ . I can also verify directly with algebra. Write  $\tilde{\mathbf{g}}_t = \begin{pmatrix} \tilde{g}_t & \tilde{g}_t^* \end{pmatrix}^T$ , where the star denotes the posterior mean for the foreign consumption distribution. Let  $M_{ij}$  denote the  $ij$  element of matrix  $M$ . Expanding equation 3.6 shows that the expected value of  $\tilde{g}_{t+1}$  under subjective beliefs is:

$$\tilde{\mathbb{E}}_t [M_{11}\Delta c_{t+1} + M_{12}\Delta c_{t+1}^*] + (1 - M_{11})\tilde{g}_t - M_{12}\tilde{g}_t^* \quad (3.9)$$

$$= \tilde{g}_t + M_{11}(\tilde{\mathbb{E}}_t[\Delta c_{t+1}] - \tilde{g}_t) + M_{12}(\tilde{\mathbb{E}}_t[\Delta c_{t+1}^*] - \tilde{g}_t^*) \quad (3.10)$$

$$= \tilde{g}_t \quad (3.11)$$

Which does not depend on  $\tilde{g}_t^*$ .

### 3.2 Model Description

I now apply the new learning process to a model of priced parameter uncertainty. Consumption is again independent across countries and time. Agents now have Epstein-Zin preferences, and so the log SDF is:

$$\log m_{t+1} = \theta \log \delta - \gamma \Delta c_{t+1} + (\theta - 1) \log \frac{PC_{t+1} + 1}{PC_t} \quad (3.12)$$

where  $\psi$  is the elasticity of intertemporal substitution,  $\gamma$  is relative risk aversion over consumption growth,  $\theta = \frac{1-\gamma}{1-\psi^{-1}}$ , and  $PC_t$  is the price to consumption ratio at time  $t$ .

Again, agents price assets using the subjective Euler equation. Beliefs evolve according to the update equations 3.6. Agents incorporate parameter uncertainty into prices. They are aware that their belief about consumption growth may change in the future, and price this risk. [Collin-Dufresne et al. \(2015\)](#) show that incorporating priced consumption risk can substantially raise Sharpe ratios when agents have Epstein-Zin preferences. Exchange rates and forward contracts are determined by complete markets and covered interest parity, just as in the power utility case. Appendix D details the numerical solution methodology.

### 3.3 Discussion of Results

I simulate the model for 40,000 quarters. I set time preference  $\delta = 0.995$  and  $\alpha = 0.02$  just as in the power utility case. I set  $\psi = 1.5$  to match parameterizations seen in the long run risks literature. I consider two cases for risk aversion:  $\gamma = 5$  and  $\gamma = 10$ , as well as three possibilities for the correlation between beliefs. I choose to model zero correlation to match the power utility model, a perfect correlation in order to see how much traction the model can get on exchange rate smoothness, as well as an intermediate correlation of 0.77 to match the cross-country correlation in log dividend price ratios reported in [Colacito and Croce \(2011\)](#).

Table 2 shows that Epstein-Zin preferences raise Sharpe ratios on the wealth claim. But the table also illustrates the tensions that arise from trying to match other moments.

First, observe that carry trade Sharpe ratios are always highest in the case of uncorrelated beliefs and lowest in the case of perfectly correlated beliefs. This is to be expected as the carry trade premium arises in my model because of *differences* in beliefs about consumption growth. When belief updates become correlated these differences become smaller, and so Sharpe ratios fall.

But focusing on the case of uncorrelated belief updates then makes exchange rates highly volatile. In the Epstein-Zin model, both the subjective and objective SDFs are driven by volatile price consumption

ratios, which are ultimately tied to fluctuating beliefs. Since exchange rates are determined by differences in subjective SDFs, then exchange rates are inevitably volatile. In contrast, the power utility model had a smooth subjective SDF and generated a volatile objective SDF through a belief distortion term. Since that belief distortion term does not enter into the exchange rate, the power utility model was able to have large carry trade Sharpe ratios alongside smooth exchange rates.

There is also a tension between exchange rate smoothness and large equity Sharpe ratios. If one would like to get close to the large equity Sharpe ratios in the data, one would need to consider a risk aversion  $\gamma$  of at least 10. But in those cases, the exchange rate is volatile even when the two countries have perfectly correlated beliefs. The exchange rate volatility rises entirely from the consumption growth portion of the Epstein-Zin SDF, and is sufficient to result in exchange rate volatility twice of that which we see in the data.

Thus the exchange rate smoothness puzzle for models of subjective beliefs can be restated as:

Large equity Sharpe ratios tell us that beliefs must be highly volatile. But given the low correlation in beliefs implied by large carry trade Sharpe ratios, exchange rates are puzzlingly smooth.

The logic behind this statement can best be captured through re-interpreting the exchange rate smoothness bound 2.29 as a set of moment restrictions on the subjective SDFs. This is possible because the relationship between the exchange rate and SDFs applies to both subjective and objective SDFs. If I allow beliefs to be uncorrelated, then the subjective SDFs are independent. In that case the variance in the exchange rate is equal to the sums of the subjective SDF variances. If I insist that beliefs are unconditionally unbiased, then to get large wealth Sharpe ratios the subjective SDF must be highly volatile. But that means exchange rates are no longer smooth.

## 4 Empirical Test: Growth Carry

My models predict that sorting currencies by exponential moving averages of past consumption growth rates should deliver comparable returns to the standard carry trade that sorts based on interest rates. I investigate this prediction for a panel of 14 developed world countries. I focus on developed world countries because my theory only applies to countries with relatively similar growth rates that do not vary substantially over time. Appendix F contains further details of the data used for this analysis. I initialize each country's growth rate expectation at 2.32%, the pre-1980 average per capita consumption

growth across all the developed countries in my sample.<sup>16</sup> Then for each country, I recursively generate beliefs according to the constant gain learning formula 2.3. This generates a sequence of growth rate beliefs  $\tilde{g}_{i,t}$  for country  $i = 1, \dots, N$  at time  $t$ . I also compute forward discounts  $fd_{i,t} = f_{i,t \rightarrow t+1} - s_t$  that correspond to interest rate differentials.

I define two kinds of carry trades. First is the standard interest rate carry, which enters into positions  $fd_{i,t} - \frac{1}{N} \sum_{i=1}^N fd_{i,t}$  in each currency  $i$  at time  $t$ . This strategy takes larger positions in currencies that have high forward discounts relative to other currencies in the cross section, and corresponds to a continuous version of the portfolio strategy specified in Lustig and Verdelhan (2007). Second, I define a growth carry, which enters into positions  $\tilde{g}_{i,t} - \frac{1}{N} \sum_{i=1}^N \tilde{g}_{i,t}$  in each currency  $i$  at time  $t$ . This strategy takes larger positions in currencies of countries that have high expected growth rates relative to other currencies in the cross section. If true growth rates were constant and growth rate expectations evolved exactly according to my constant gain specification, then these sorts would be identical.

Figure 2 shows that these two strategies share similar periods of gains and losses. This suggests that high interest rate currencies correspond to those with high expected growth rates. While the growth carry earns a Sharpe ratio of 0.26, the interest rate carry earns a Sharpe ratio of 0.40. Given that estimates based on macro data are inevitably noisier, and given that my constant gain learning process inevitably suffers from misspecification, these the comparable Sharpe ratios indicate that differences in growth rate expectations do a good job in explaining the dispersion of forward discounts and carry trade returns.

This growth carry is a novel empirical test that helps to distinguish competing stories of the carry trade. For example, the long run risks model of Bansal and Shaliastovich (2012) and the rare disasters model of Farhi and Gabaix (2014) do not feature expected returns that depend on long moving averages of past consumption growth. In fact, in the domestic asset pricing literature, it is an advantage of the long run risks model that past consumption growth does not forecast the state variable (the price to consumption ratio) that forecasts expected returns. But this strength of the long run risks model in the domestic realm is a weakness in resolving the carry trade. In the case of the carry trade, the state variable that forecasts expected returns, the interest rate differential, is predictable on the basis of past consumption growth. Therefore the growth carry is support for my beliefs based story as well as the habits based story of Verdelhan (2010).

In table 3, I provide a summary of the empirical performance of rational expectations solutions to the carry trade as well as a subset of my belief based stories. Overall, the beliefs model with power utility is the only model to match all of the foreign exchange facts: large carry trade returns, smooth exchange

---

<sup>16</sup>The exact level of this belief is not very important because the carry trades that I study are all based on the cross section of growth rates.

rates, and a growth carry. However, it fails to match a large equity Sharpe ratio. When one imposes the requirement that models must also match large equity Sharpe ratio, then the model of beliefs that matches the most facts is the Epstein-Zin model with high risk aversion  $\gamma = 10$ . However, in that case exchange rates are no longer smooth, and my model with Epstein-Zin preferences and priced parameter uncertainty matches the same set of facts as the habits model.

## 5 Conclusion

I extend the modern asset pricing literature on parameter learning to the open economy and show how errors in beliefs about mean consumption growth can provide a plausible account of persistent and large carry trade premia. In my model, the carry trade is profitable not because of differences in conditional SDF variances, but rather because of differences in beliefs about mean consumption growth in different countries. I also show how incorporating subjective beliefs in open economy asset pricing models can explain how smooth exchange rates are consistent with large carry trade Sharpe ratios. I then extend the model to the case of Epstein-Zin preferences, and find that my simple model of international learning cannot jointly match high equity Sharpe ratios, high carry trade Sharpe ratios, and smooth exchange rates. My beliefs based stories predict that a growth carry that sorts countries based on expected growth rates instead of interest rates should earn returns comparable to the interest rate carry, and I find support for this in the data. I show how this test distinguishes my model relative to existing rational expectations models.

A key extension of my model would be to incorporate how beliefs may interact with incomplete markets. Given that learning seems to have traction even in simple consumption environments, one could extend the rational expectations analysis of [Lustig and Verdelhan \(2016\)](#) to see how incomplete market wedges combined with beliefs can help match key moments of international asset returns. Another direction would be to find ways to model the non-linearities in carry trade returns. While existing explanations of the skewness of currency returns focus on intermediary constraints ([Brunnermeier et al., 2008](#)), it is an open question if and how learning might generate similar results.

## References

- Abel, Andrew**, "An exploration of the effects of pessimism and doubt on asset returns," *Journal of Economic Dynamics and Control*, 2002, 26.
- Backus, David K, Silverio Foresi, and Chris I Telmer**, "Affine Term Structure Models and the Forward Premium Anomaly," *Journal of Finance*, February 2001, 56 (1).
- Bansal, Ravi and Ivan Shaliastovich**, "A Long-Run Risks Explanation of Predictability Puzzles in Bond and Currency Markets," *Review of Financial Studies*, 2012, 26 (1), 1–33.
- Brandt, Michael W., John H. Cochrane, and Pedro Santa-Clara**, "International risk sharing is better than you think, or exchange rates are too smooth," *Journal of Monetary Economics*, 2006, 53, 671–698.
- Brunnermeier, Markus K., Lasse Pedersen, and Stefan Nagel**, *Carry Trades and Currency Crashes*, Vol. 23 of *NBER Macroeconomics Annual*, University of Chicago Press,
- Cogley, Timothy and Thomas J Sargent**, "Market price of risk and the equity premium: A legacy of the Great Depression," *Journal of Monetary Economics*, 2008, 55, 454–476.
- Colacito, Riccardo and Mariano M. Croce**, "Risks for the Long Run and the Real Exchange Rate," *Journal of Political Economy*, 2011, 119 (1), 153–182.
- Collin-Dufresne, Pierre, Michael Johannes, and Lars A. Lochstoer**, "Parameter Learning in General Equilibrium: The Asset Pricing Implications," *Working Paper*, 2015.
- Farhi, Emmanuel and Xavier Gabaix**, "Rare Disasters and Exchange Rates," *Working Paper*, October 2014.
- Hansen, Lars Peter and Ravi Jagannathan**, "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy*, April 1991, 99 (2), 225–262.
- Lewellen, Jonathan and Jay Shanken**, "Asset-Pricing Tests, and Market Efficiency," *Journal of Finance*, June 2002, 57 (3), 1113–1145.
- Lustig, Hanno and Adrien Verdelhan**, "The Cross Section of Foreign Currency Risk Premia and Consumption Growth Risk," *American Economic Review*, March 2007.
- **and –**, "Comment on "Carry Trades and Currency Crashes"," in Daron Acemoglu, Kenneth Rogoff, and Michael Woodford, eds., *NBER Macroeconomics Annual 2008*, Vol. 23 2008, pp. 361–384.
- **and –**, "Does Incomplete Spanning in International Financial Markets Help to Explain Exchange Rates?," *NBER Working Paper*, February 2016.
- **, Nikolai Roussanov, and Adrien Verdelhan**, "Common Risk Factors in Currency Markets," *The Review of Financial Studies*, November 2011, 24 (11), 3731–3777.
- Malmendier, Ulrike and Stefan Nagel**, "Learning from Inflation Experiences," *Quarterly Journal of Economics*, October 2015.
- Milani, Fabio**, "Expectations, learning and macroeconomic persistence," *Journal of Monetary Economics*, January 2007, 54, 2065–2082.
- Pastor, Lubos and Pietro Veronesi**, "Learning in Financial Markets," *Annual Review of Financial Economics*, 2009.
- Verdelhan, Adrien**, "A Habit-Based Explanation of the Exchange Rate Premium," *Journal of Finance*, February 2010, 65 (1), 123–146.

## A Tables

**Table 1:** Moments from model simulations of the power utility model. I run quarterly simulations but annualize all moments. The subjective SDF refers to the experienced marginal utility growth of the agent. The objective SDF refers to the subjective SDF plus the common belief distortion. The carry trade refers to the strategy of buying the high interest rate currency and shorting the low interest rate currency. The scaled carry trade scales the carry trade exposure according to the size of the forward discount  $r_t^{f,*} - r_t^f$ , and therefore takes advantage of the higher conditional Sharpe ratio when the forward discount is large.

Moments			
Consumption	Consumption Growth (%)	$g$	1.51
	Consumption Volatility (%)	$\sigma(\Delta c)$	1.67
Returns + Beliefs	Carry Trade SR	$\text{SR}([f_{t \rightarrow t+1} - s_{t+1}] \times \text{sign}[r_t^{f,*} - r_t^f])$	0.16
	Scaled Carry Trade SR	$\text{SR}([f_{t \rightarrow t+1} - s_{t+1}] \times [r_t^{f,*} - r_t^f])$	0.21
	Wealth SR	$\text{SR}(r_W)$	0.05
	Volatility of Belief (%)	$\sigma(\tilde{g}_t)$	0.16
SDF Moments	Subjective SDF Volatility	$\sigma(m)$	0.08
	Objective SDF Volatility	$\sigma(\hat{m})$	0.30
	Subjective SDF Correlation	$\text{Corr}(m, m^*)$	0.00
	Objective SDF Correlation	$\text{Corr}(\hat{m}, \hat{m})$	0.92
	Exchange Rate Volatility	$\sigma(\Delta s)$	0.12

**Table 2:** Moments from model simulations of the Epstein-Zin model with priced parameter uncertainty. I run quarterly simulations but annualize all moments. The carry trade refers to the strategy of buying the high interest rate currency and shorting the low interest rate currency. The scaled carry trade scales the carry trade exposure according to the size of the forward discount  $r_t^{f,*} - r_t^f$ , and therefore takes advantage of the higher conditional Sharpe ratio when the forward discount is large.

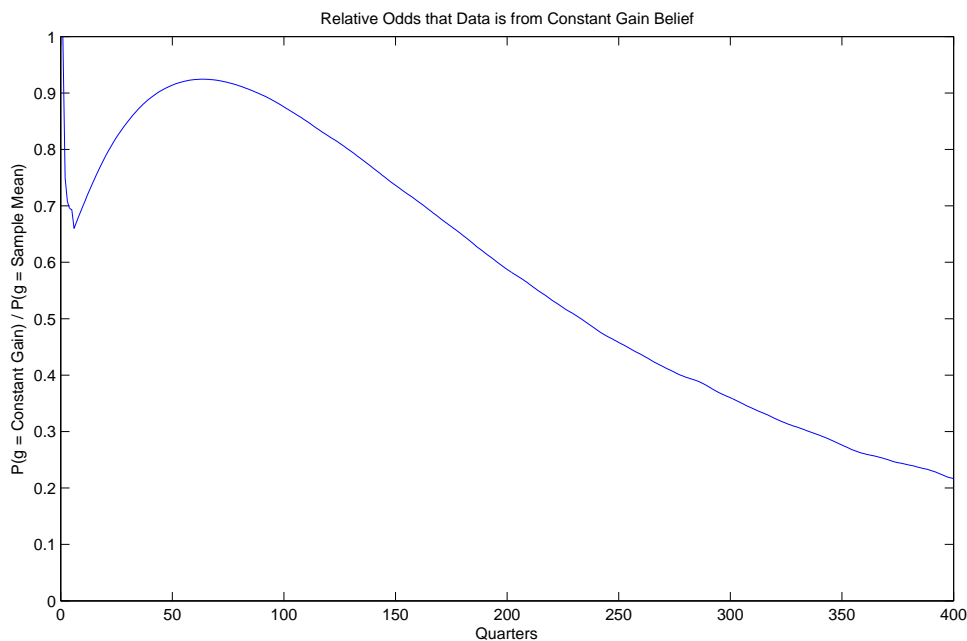
Risk Aversion Belief Correlation	$\gamma = 5$			$\gamma = 10$		
	$\rho = 0$	$\rho = 0.77$	$\rho = 1$	$\rho = 0$	$\rho = 0.77$	$\rho = 1$
Carry Trade SR	0.14	0.04	0.00	0.13	0.06	0.00
Scaled Carry SR	0.18	0.06	0.00	0.17	0.07	0.00
Wealth SR	0.19	0.13	0.13	0.36	0.28	0.27
Exchange Rate Volatility	0.44	0.17	0.12	0.76	0.32	0.24



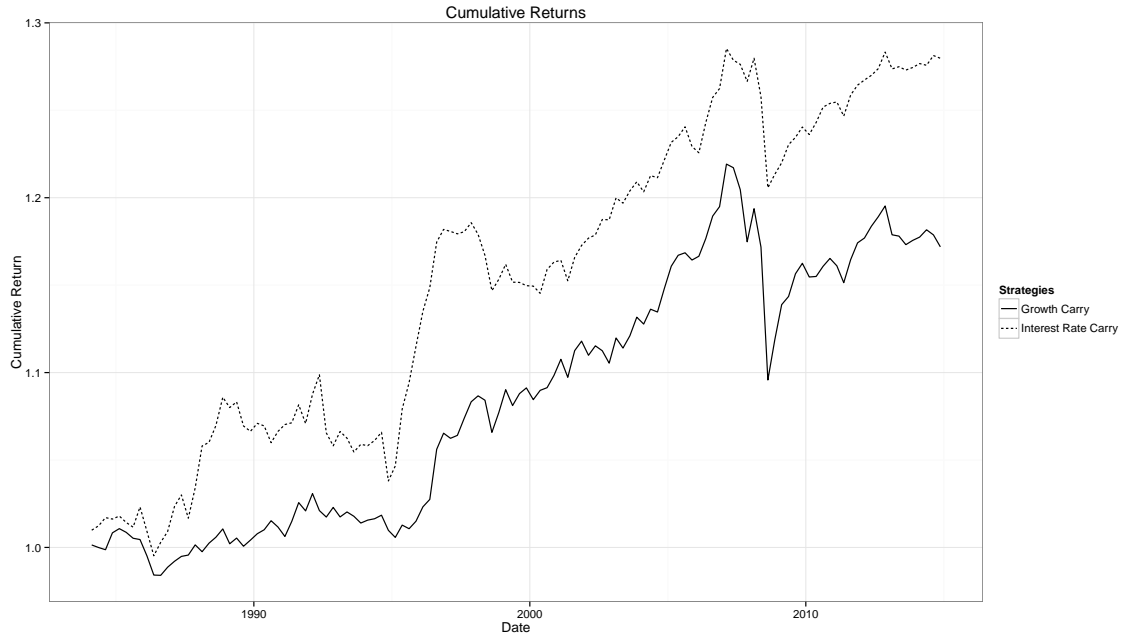
**Table 3:** Summary of how models perform on replicating different features of exchange rates and the carry trade in the data. Carry SR refers to the ability of a model to generate large carry trade SR. Smooth FX refers to the ability of the model to limit the standard deviation of the real exchange rate to around 0.10. Growth carry refers to the ability of the model to link the cross sectional dispersion in interest rates to recent levels of consumption growth. Wealth SR refers to the ability of a model to generate equity Sharpe ratios near 0.40. The ✓'s mark the dimensions on which each model succeeds.

Data Feature	Habits (Verdelhan, 2010)	Long Run Risks (Bansal and Shaliastovich, 2012)	Rare Disasters (Farhi and Gabaix, 2014)	Beliefs + Power Utility, $\gamma = 5$	Beliefs + EZ + Independent Beliefs, $\gamma = 10$
Carry SR	✓	✓	✓	✓	✓
Smooth FX		✓	✓	✓	
Growth Carry	✓			✓	✓
Wealth SR	✓	✓	✓		✓

## B Figures



**Figure 1:** Average relative odds of data being generated from the constant gain belief  $\tilde{g}_t$  instead of the sample mean  $\bar{g}_t$ , for  $t = 1, \dots, T = 400$  quarters. In each simulation trial  $i = 1, \dots, N = 10,000$ , I simulate a sequence of consumption growth  $\Delta c_{i,t}$ . Constant gain beliefs  $\tilde{g}_{t,i}$  are initialized with  $\tilde{g}_{0,i} = g$  and evolve according to  $\tilde{g}_{t,i} = \alpha \Delta c_{t,i} + (1 - \alpha) \tilde{g}_{t-1,i}$ . The full Bayesian belief is the sample mean  $\tilde{g}_{t,i} = \frac{1}{t} \sum_{j=1}^t \Delta c_{j,i}$ . In each simulation trial, I compute the log likelihood ratio  $\ell_{t,i} = -\frac{1}{2\sigma^2} \sum_{j=1}^t \left[ (\Delta c_{j,i} - \tilde{g}_{t,i})^2 - (\Delta c_{j,i} - \bar{g}_{t,i})^2 \right]$ . The curve above, at quarter  $t$ , is then  $\exp\left(\frac{1}{N} \sum_{i=1}^N \ell_{t,i}\right)$ . The quantity represents the relative likelihood of the data being generated from a process of consumption growth with the mean  $\tilde{g}_t$  versus being generated from  $\bar{g}_t$ . Given that the sample mean is the maximum likelihood estimate, this relative likelihood is always less than one, although it does decrease only slowly.



**Figure 2:** Comparison of carry trade strategies on developed world countries. The growth carry refers to the long short strategy sorting on model implied consumption growth beliefs, entering in positions  $\left(\tilde{g}_{i,t} - \frac{1}{N} \sum_{i=1}^N \tilde{g}_{i,t}\right)$  in currency  $i$  at time  $t$ , when the growth rate belief for country  $i$  at time  $t$  is  $\tilde{g}_{i,t}$ . The interest rate carry refers to the long short strategy sorting on forward discounts, entering in positions  $\left(fd_{i,t} - \frac{1}{N} \sum_{i=1}^N fd_{i,t}\right)$  in currency  $i$  at time  $t$ , when the forward discount is defined as the log difference between the forward contract and the spot  $fd_{i,t} = f_{i,t \rightarrow t+1} - s_{i,t}$ . Log returns to each strategy are scaled to have a standard deviation of 1%.

## C Proofs of Multivariate Constant Gain Learning

Let consumption growth follow

$$\Delta \mathbf{c}_{t+1} = \begin{pmatrix} \Delta c_{t+1} \\ \Delta c_{t+1}^* \end{pmatrix} \sim \text{MVN}(\mathbf{g}, \Sigma), \mathbf{g} = \begin{pmatrix} g \\ g^* \end{pmatrix}, \Sigma = \sigma^2 I$$

Impose the prior

$$p(g, g^*) \sim \text{MVN}(\mathbf{g}, Y), \mathbf{g} = \begin{pmatrix} g \\ g^* \end{pmatrix}, Y = \tau^{-2} \begin{pmatrix} 1 & \phi \\ \phi & 1 \end{pmatrix}$$

I start with the full Bayesian solution. After  $t$  observations, the posterior for the means is a multivariate normal with mean  $\tilde{\mathbf{g}}_t$

$$\tilde{\mathbf{g}}_t = \left( Y^{-1} + t\Sigma^{-1} \right)^{-1} \left( Y^{-1}\mathbf{g} + \Sigma^{-1} \sum_{s=1}^t \Delta \mathbf{c}_s \right)$$

This can be written recursively as:

$$\begin{aligned} \tilde{\mathbf{g}}_t &= \left( \Sigma Y^{-1} + tI \right)^{-1} \Delta \mathbf{c}_t + \left( I - \left( \Sigma Y^{-1} + tI \right)^{-1} \right) \tilde{\mathbf{g}}_{t-1} \\ \tilde{\mathbf{g}}_0 &= \mathbf{g} \end{aligned}$$

The first step to get to constant gain learning is to replace the true sample size  $t$  with the effective sample size  $\alpha^{-1}$ . After doing so, expand out the multiplier matrix on consumption growth:

$$\left( \Sigma Y^{-1} + \alpha^{-1} I \right)^{-1} = \begin{pmatrix} \alpha^{-1}(1 - \phi^2) + \sigma^2 \tau^2 & \phi \sigma^2 \tau^2 \\ \phi \sigma^2 \tau^2 & \alpha^{-1}(1 - \phi^2) + \sigma^2 \tau^2 \end{pmatrix} \times \frac{1}{\alpha^{-2}(1 - \phi^2) + 2\alpha^{-1}\sigma^2 \tau^2 + \sigma^2 \tau^4}$$

The general strategy involves sending  $\tau \rightarrow 0$  and seeing what the resulting matrix implies for learning. First consider the special case of  $\phi = 1$ . In that case I get the symmetric matrix

$$\frac{\sigma^2 \tau^2}{2\alpha^{-1}\sigma^2 \tau^2 + \sigma^2 \tau^4} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Let the precision  $\tau \rightarrow 0$  to reflect a diffuse prior. The matrix then becomes

$$\frac{\alpha}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Concretely, what this means is that  $\tilde{g}_t$  evolves according to

$$\tilde{g}_t = \alpha \cdot \frac{\Delta c_t + \Delta c_t^*}{2} + (1 - \alpha) \tilde{g}_{t-1}$$

Therefore the two countries' beliefs are perfectly correlated because the new information incorporated at each stage  $t$  is just the average of the two countries' consumption growth at time  $t$ .

The case of  $\phi = 0$  also accords with the intuition that independent priors implies independent learning. In that case,

$$\begin{aligned} (\Sigma Y^{-1} + \alpha^{-1} I)^{-1} &= \begin{pmatrix} \alpha^{-1} + \sigma^2 \tau^2 & \\ & \alpha^{-1} + \sigma^2 \tau^2 \end{pmatrix} \times \frac{1}{\alpha^{-2} + 2\alpha^{-1} \sigma^2 \tau^2 + \sigma^2 \tau^4} \\ &= I \times \frac{\alpha^{-1} + \sigma^2 \tau^2}{\alpha^{-2} + 2\alpha^{-1} \sigma^2 \tau^2 + \sigma^2 \tau^4} \\ &\rightarrow I \times \frac{\alpha^{-1}}{\alpha^{-2}} = \alpha I \end{aligned}$$

Which corresponds exactly to the intuition of just using each country's past history of consumption growth.

However, letting  $\tau \rightarrow 0$  for intermediate levels of  $\phi$  does not lead to intermediate solutions. Intuitively, as  $\tau \rightarrow 0$ , the effect of  $Y$  in the updating equation goes to zero, and so the influence of the independent consumption growth overwhelms any effect of the prior in inducing correlations between beliefs. Formally, one can take limits of the updating matrix  $(\Sigma Y^{-1} + \alpha^{-1} I)^{-1}$  as  $\tau \rightarrow 0$  to get:

$$\begin{aligned} (\Sigma Y^{-1} + \alpha^{-1} I)^{-1} &\rightarrow \begin{pmatrix} \alpha^{-1} (1 - \phi^2) & \\ & \alpha^{-1} (1 - \phi^2) \end{pmatrix} \times \frac{1}{\alpha^{-2} (1 - \phi^2)} \\ &= \alpha I \end{aligned}$$

Therefore to get arbitrarily correlated beliefs under diffuse priors, I model  $\phi$  with local to unity asymp-

otics. Let  $\phi \sim 1 - c\tau^2 + o(\tau^2)$ , and therefore  $\phi^2 \sim 1 - 2c\tau^2 + o(\tau^2)$ . In that case as  $\tau \rightarrow 0$

$$\begin{aligned}
& \left( \Sigma Y^{-1} + \alpha^{-1} I \right)^{-1} \\
= & \begin{pmatrix} \alpha^{-1}(1 - \phi^2) + \sigma^2 \tau^2 & \phi \sigma^2 \tau^2 \\ \phi \sigma^2 \tau^2 & \alpha^{-1}(1 - \phi^2) + \sigma^2 \tau^2 \end{pmatrix} \times \frac{1}{\alpha^{-2}(1 - \phi^2) + 2\alpha^{-1}\sigma^2 \tau^2 + \sigma^2 \tau^4} \\
\sim & \begin{pmatrix} \alpha^{-1} 2c\tau^2 + \sigma^2 \tau^2 & (1 - c\tau^2) \sigma^2 \tau^2 \\ (1 - c\tau^2) \sigma^2 \tau^2 & \alpha^{-1} 2c\tau^2 + \sigma^2 \tau^2 \end{pmatrix} \times \frac{1}{\alpha^{-2} 2c\tau^2 + 2\alpha^{-1}\sigma^2 \tau^2 + \sigma^2 \tau^4} \\
\rightarrow & \begin{pmatrix} 2\alpha^{-1}c + \sigma^2 & \sigma^2 \\ \sigma^2 & 2\alpha^{-1}c + \sigma^2 \end{pmatrix} \times \frac{1}{2c\alpha^{-2} + 2\alpha^{-1}\sigma^2} \\
= & \alpha \begin{pmatrix} 2\alpha^{-1}c + \sigma^2 & \sigma^2 \\ \sigma^2 & 2\alpha^{-1}c + \sigma^2 \end{pmatrix} \times \frac{1}{2c\alpha^{-1} + 2\sigma^2} \\
= & \begin{pmatrix} \frac{\sigma^2}{2c} + \alpha^{-1} & -\frac{\sigma^2}{2c} \\ -\frac{\sigma^2}{2c} & \frac{\sigma^2}{2c} + \alpha^{-1} \end{pmatrix}^{-1} \\
= & \left( \frac{\sigma^2}{2c} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \alpha^{-1} I \right)^{-1}
\end{aligned}$$

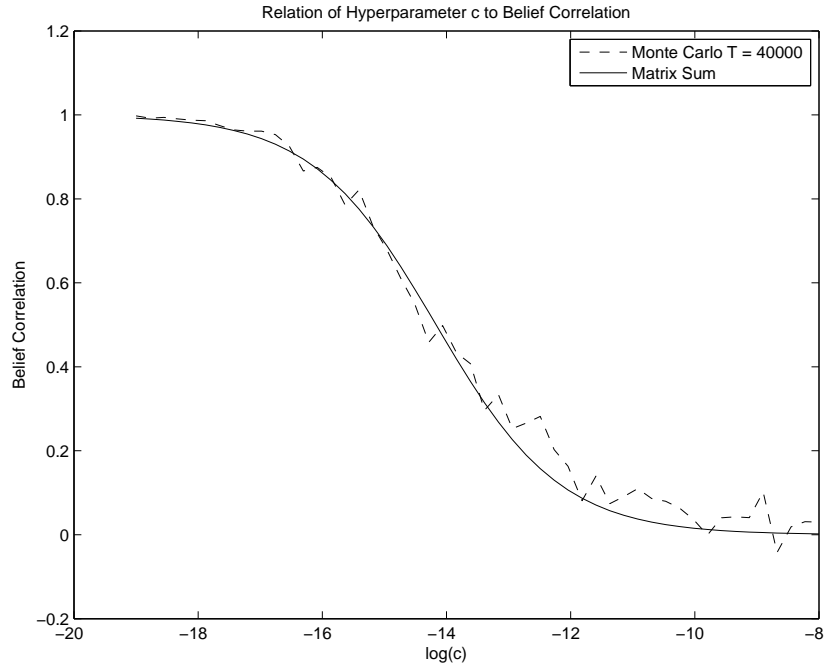
In deriving this learning model I imposed a hyperparameter  $c$ . One way to calibrate  $c$  is to think about the implied correlation between the two beliefs  $\tilde{g}_t, \tilde{g}_t^*$ . Let  $M = \left( \frac{\sigma^2}{2c} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \alpha^{-1} I \right)^{-1}$ . Because  $M$  is the inverse of a symmetric matrix, it too is symmetric. Iterating the recursion for  $\tilde{\mathbf{g}}_t$  backwards yields:

$$\begin{aligned}
\tilde{\mathbf{g}}_t &= M\Delta\mathbf{c}_n + (I - M)\tilde{\mathbf{g}}_{t-1} \\
&= M\Delta\mathbf{c}_n + (I - M)[M\Delta\mathbf{c}_{n-1} + (I - M)\tilde{\theta}_{n-2}] \\
&= \sum_{j=0}^{\infty} (I - M)^j M\Delta\mathbf{c}_{n-j}
\end{aligned}$$

Note that  $M, I - M$  commute, and that  $M = M^T$ . Therefore

$$\begin{aligned} \text{Var}(\tilde{\mathbf{g}}_i) &= \sum_{j=0}^{\infty} (I - M)^j M \Sigma M^T (I - M^T)^j \\ &= \sigma^2 \sum_{j=0}^{\infty} (I - M)^j M^2 (I - M)^j \\ &= \sigma^2 M^2 \sum_{j=0}^{\infty} (I - M)^{2j} \end{aligned}$$

This provides both the parameter uncertainty matrix and as well as the correlation between beliefs. Figure 3 shows the relationship between the hyperparameter and the belief correlation, with the belief correlation both computed analytically and checked by simulation.



**Figure 3:** Relationship of hyperparameter  $c$  to the correlation of beliefs. Monte Carlo refers to the correlation as computed from a Monte Carlo simulation of how the beliefs evolve. Matrix sum refers to explicitly computing the infinite sum in the matrix expression.

## D Solving for the Price Consumption Ratio in the Epstein-Zin Case

Start with the Euler equation for the return on wealth

$$1 = \tilde{\mathbb{E}} \left[ M_{t+1} R_{t+1}^W \right]$$

Now note that the return on wealth can be expressed as

$$R_{t+1}^W = \frac{PC_{t+1} + 1}{PC_t} \frac{C_{t+1}}{C_t}$$

Plugging this into the Epstein-Zin SDF 3.12 yields

$$\begin{aligned} 1 &= \tilde{\mathbb{E}} \left[ \delta^\theta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma+1} \left( \frac{PC_{t+1} + 1}{PC_t} \right)^\theta \right] \\ PC_t &= \tilde{\mathbb{E}} \left[ \delta^\theta \exp((-\gamma + 1) \Delta c_{t+1}) (PC_{t+1} + 1)^\theta \right]^{1/\theta} \end{aligned}$$

Since agents price parameter uncertainty, they understand that  $PC_{t+1}$  will differ from today as a function of what the subjective expected consumption growth rate will be tomorrow. Translate the expectation into an integral over the subjective distribution  $p(\Delta \mathbf{c}_{t+1}; \hat{\mathbf{g}}_t, \tilde{\Sigma})$ , where  $p(s_1, s_2; \boldsymbol{\mu}, \Sigma)$  is the density of a multivariate normal with mean  $\boldsymbol{\mu}$  and covariance  $\Sigma$ ,  $\hat{\mathbf{g}}_t$  is the time  $t$  belief about expected future consumption growth, and  $\tilde{\Sigma} = \Sigma + \text{Var}(\hat{\mathbf{g}}_t)$  captures both parameter uncertainty (computed in Appendix C) and the physical variance of consumption growth.

I show in the text that the pricing of the consumption claim is independent of beliefs about foreign consumption. Define  $\mathbf{s} = \begin{pmatrix} s_1 & s_2 \end{pmatrix}_T$ ,  $\hat{\mathbf{g}}_t = \begin{pmatrix} \tilde{g}_t & 0 \end{pmatrix}_T$ . Let  $M$  be the multiplier matrix on consumption growth in the belief formation process described in equation 3.6 and let  $M_{ij}$  be the  $ij$  element of that matrix. Then the price consumption ratio, as a function of the expected growth rate for domestic consumption  $\tilde{g}_t$ , can be expressed as,

$$PC(\tilde{g}_t) = \left[ \int p(s_1, s_2; \hat{\mathbf{g}}_t, \tilde{\Sigma}) \delta^\theta \cdot e^{(-\gamma+1)s_1} \cdot [PC(M_{11}s_1 + (1 - M_{11})\tilde{g}_t) + 1]^\theta ds_1 ds_2 \right]^{1/\theta} \quad (\text{D.1})$$

Where  $s_1, s_2$  are dummy variables for consumption growth. To get the risk free rate, note that  $R_f = \tilde{\mathbb{E}} [\tilde{M}_{t+1}]^{-1}$ . Therefore



$$R_f(\tilde{g}_t) = \left[ \int p(s_1, s_2; \hat{\mathbf{g}}_t, \tilde{\Sigma}) \delta^\theta \cdot e^{-\gamma s_1} \cdot \left[ \frac{PC(M_{11}s_1 + (1 - M_{11})\tilde{g}_t) + 1}{PC(\tilde{g}_t)} \right]^{\theta-1} ds_1 ds_2 \right]^{-1} \quad (\text{D.2})$$

To search for the proper function  $PC(\tilde{g})$ , I use cubic interpolation for  $\log PC(\tilde{g})$ . That is, I assume  $\log PC(\tilde{g})$  is locally cubic in  $\tilde{g}$ . Initialize a grid  $G$  for  $\tilde{g}$  of 100 linearly spaced points between  $-0.01/4$  and  $0.05/4$ . It's important to go far beyond plausible because truncation at the boundaries distort the distribution of returns.<sup>17</sup> Set  $F_0 = 0$ . Then consider the functional recursion

$$F_{k+1}(\tilde{g}_t) = \frac{1}{\theta} \log \int p(s_1, s_2; \hat{\mathbf{g}}_t, \tilde{\Sigma}) \delta^\theta \cdot \exp((-\gamma + 1)s_1) \cdot [\exp(F_k(M_{11}s_1 + (1 - M_{11})\tilde{g}_t)) + 1]^\theta ds_1 ds_2$$

Where  $F_k$  is evaluated using cubic interpolation over the grid  $G$ . If an input to  $F_k$  would go outside the grid, use the nearest  $g \in G$  instead.

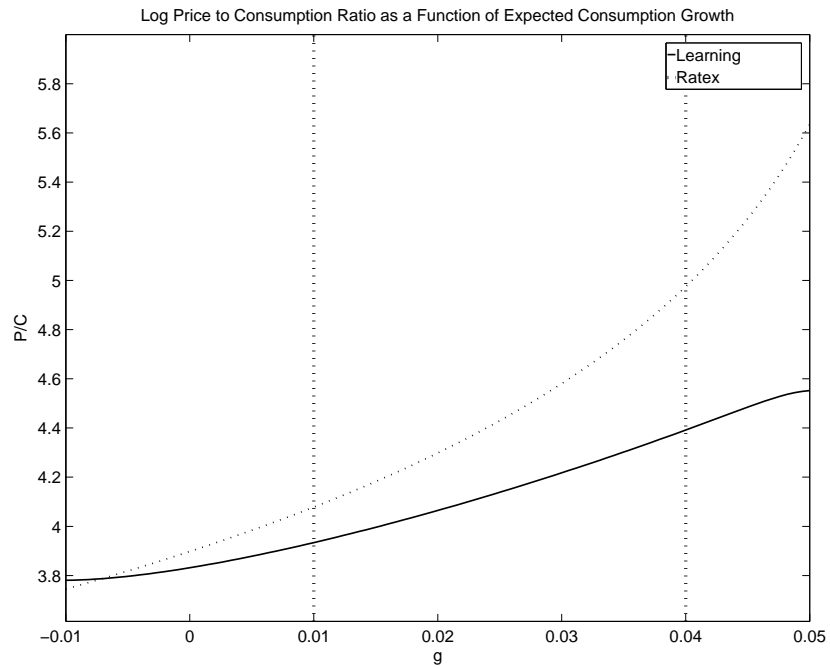
At each step, define  $\Delta_{k+1} = \max_{g \in G} |F_{k+1}(g) - F_k(g)|$ . I want to converge until I am reasonably confident that there is little distance left for each of the values to move. Since the  $\Delta_{k+1}$  decay slowly, it will be important to take into account the full decay. Therefore define the ratio  $\rho_{k+1} = \Delta_{k+1}/\Delta_k$ . In my simulations  $\rho_{k+1} < 1$ , with a value near 0.996. Then define  $\varepsilon_{k+1} = \frac{\Delta_{k+1}}{1-\rho_{k+1}}$ . The idea here is that  $\varepsilon_{k+1}$  serves as an estimate on the further change in values of  $F$  assuming the changes keep on decaying at rate  $\rho_{k+1}$ . Stop the convergence when  $\varepsilon_{k+1} < \varepsilon_{max}$ . In my simulations I set  $\varepsilon_{max} = 0.005$ , so that the price to consumption ratios are off by no more than 1%. Once this process converges to a function  $F$ , let  $PC(\tilde{g}) = \exp(F(\tilde{g}))$ . With this  $PC(\tilde{g})$  function, I can also then compute the risk free rate function  $R(\tilde{g})$  using equation D.2.

For all of the integrations discussed, I integrate over the interval  $[\tilde{g} - 5\tilde{\sigma}, \tilde{g} + 5\tilde{\sigma}]^2 \subset \mathbb{R}^2$ , where  $\tilde{\sigma}$  is the standard deviation of the subjective predictive distribution for consumption growth.

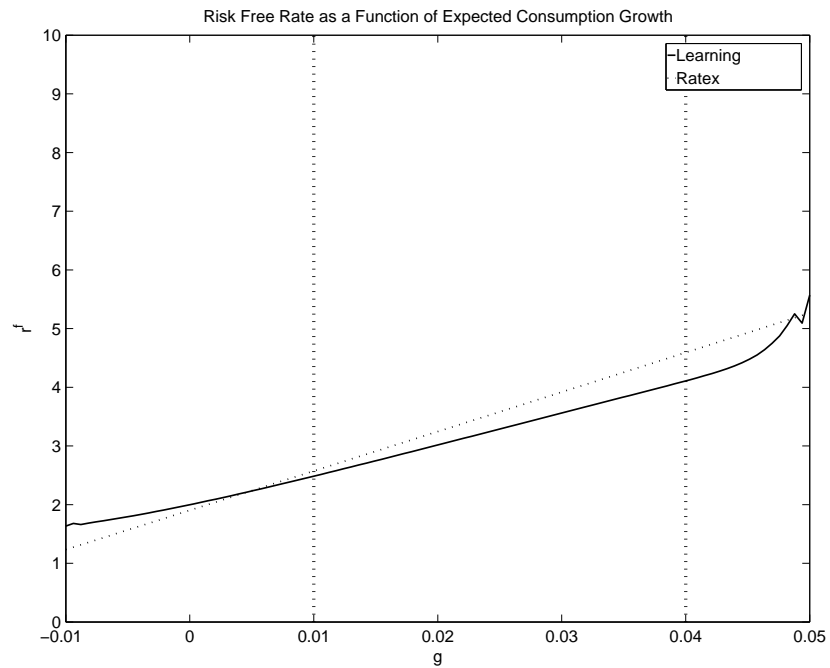
Figures 4, 5 show how  $\log PC$  and  $r^f$  relate to the subjective expected growth rate  $\tilde{g}$ , and compare these quantities to the values when the parameters are known with certainty. These are simulations under  $\gamma = 5$  and with the correlation between beliefs in the two countries set at 0. In general, the price to consumption ratio and risk free rate are increasing in the growth rate.

---

<sup>17</sup>For example, when you are at the bottom endpoint, then agents believe that expected consumption growth can only rise. This imparts a positive bias to the wealth return and, because of Epstein-Zin preferences, distorts the risk free rate as well.



**Figure 4:** Annualized log price to consumption ratio under parameter learning and rational expectations. Vertical dotted lines mark the points where the annual expected consumption growth is 1 and 4%.



**Figure 5:** Annualized log risk free rate in percent under parameter learning and rational expectations. Vertical dotted lines mark the points where the annual expected consumption growth is 1 and 4%.

## E Simulation Methodology

For  $t = 1$ , initialize  $\tilde{\mathbf{g}}_t = \begin{pmatrix} g & g \end{pmatrix}^T$ . Since both models provide formulas relating price to consumption and the risk free rate to subjective expected consumption growth, use  $R_f(\tilde{g}_t), PC(\tilde{g}_t)$  to initialize the risk free rate and price to consumption ratio. Define  $\mathbf{g} = \begin{pmatrix} g & g \end{pmatrix}^T$ ,  $\tilde{\mathbf{g}} = \begin{pmatrix} \tilde{g}_t & \tilde{g}_t^* \end{pmatrix}^T$ , and  $\Sigma = \sigma^2 \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$ . Define the subjective variance matrices  $\tilde{\Sigma} = \Sigma + \text{Var}(\tilde{\mathbf{g}}_t)$ , where  $\text{Var}(\tilde{\mathbf{g}}_t)$  is computed by equation 2.5 for the power utility case and computed from the analytical formulas in Appendix C for the Epstein-Zin case. Define  $\tilde{\sigma}^2$  as the 1, 1 entry of  $\text{Var}(\tilde{\mathbf{g}}_t)$ .

Then for stages  $t = 2, \dots, T$ :

Step	Description	Power Utility	Epstein-Zin
1	Draw consumption shocks		$\begin{pmatrix} \Delta c_t & \Delta c_t^* \end{pmatrix}^T \sim N(\mathbf{g}, \Sigma)$
2	Update beliefs	$\tilde{g}_t = (1 - \alpha)\tilde{g}_{t-1} + \Delta c_{t+1}$ $\tilde{g}_t^* = (1 - \alpha)\tilde{g}_{t-1}^* + \Delta c_{t+1}^*$	Follow equations 3.6
3	Get PC ratios, risk free rates	$k = \delta \exp\left((1 - \gamma)\tilde{g}_t + \frac{1}{2}(1 - \gamma)^2\tilde{\sigma}^2\right)$ $PC_t = \frac{k}{1 - k}$ $r_t^f = -\log \delta + \gamma\tilde{g}_t - \frac{1}{2}\gamma^2\tilde{\sigma}^2$	Use the numerical solution techniques detailed in appendix D and plug in $PC(\tilde{g}_t), R(\tilde{g}_t)$ .
4	Get belief distortions	$\frac{\tilde{f}_t}{f} = \frac{p(\Delta c_t, \Delta c_t^*; \tilde{\mathbf{g}}_t, \tilde{\Sigma})}{p(\Delta c_t, \Delta c_t^*; \mathbf{g}, \Sigma)}$ , where $p(\cdot; \mu, \Sigma)$ is the pdf of a multivariate normal with mean $\mu$ and covariance matrix $\Sigma$ .	
5	Get log SDF realizations ( $m$ )	$m_{t+1} = \log \delta - \gamma\Delta c_t$	$\log m_{t+1} = \theta \log \delta - \gamma\Delta c_{t+1} + (\theta - 1) \log \frac{PC_{t+1} + 1}{PC_t}$

When equations are reported only for one country the natural generalization with foreign variables applies. Given these values for both models, the process for computing the remaining moments is model free:

$$\begin{aligned}
 \Delta s_{t+1} &= m_{t+1} - m_{t+1}^* \\
 \hat{m}_{t+1} &= m_{t+1} + \frac{\tilde{f}_{t+1}}{f} \\
 r_{t+1}^{w} &= \log \frac{PC_{t+1} + 1}{PC_t} + \Delta c_{t+1} \\
 f_{t \rightarrow t+1} - s_{t+1} &= r_t^{f,*} - r_t^f - \Delta s_{t+1}
 \end{aligned}$$

## F Data

Data for quarterly real consumption come from the OECD, accessed through FRED. Annual population data come from the World Bank's World Development Indicators. I construct quarterly per capita consumption in country  $i$  at time  $t$  by

$$\Delta c_{i,t} = c_{i,t} - c_{i,t-1} - \Delta p_{i,t}$$

Where  $\Delta p_{i,t}$  is the five year moving average of *annual* population changes, divided by 4 to get a quarterly approximation, for the five years including the consumption observation. Therefore the per capita consumption data for Australia in the second quarter of 1985 uses population data from 1980 to 1985. I do this because quarterly consumption series are not uniform across countries, and population growth is sufficiently smooth so that the moving average approximation should not induce significant errors while also minimizing the effect of idiosyncratic population measurement errors in any given year.

Exchange rate data comes from Barclays and Reuters series, accessed through Datastream.

**Table 4:** Countries used in analysis

Merged Consumption and Exchange Rate Data			
Country	First Consumption Observation	First Combined Observation	Last Combined Observation
Australia	1964.2	1984.4	2014.4
Canada	1964.2	1984.4	2014.4
Denmark	1995.2	1995.2	2014.4
Europe	1995.2	1999.2	2011.4
France	1964.2	1983.4	1998.4
Germany	1970.2	1983.4	1998.3
Italy	1981.2	1984.2	1998.3
Japan	1994.2	1994.2	2014.4
Netherlands	1988.2	1988.2	1998.3
New Zealand	1988.2	1988.2	2014.4
Norway	1978.2	1984.4	2014.4
Sweden	1964.2	1984.4	2014.4
Switzerland	1980.2	1983.4	2014.4
United Kingdom	1964.2	1983.4	2014.4

Only Consumption Data		
Country	First Consumption Observation	Last Observation
United States	1964.2	2014.4
Belgium	1995.2	2014.4