Global Disaster Risk and The Carry Trade

by

Corey Feldman

An honors thesis submitted in partial fulfillment

of the requirements for the degree of

Bachelor of Science

Undergraduate College

Leonard N. Stern School of Business

New York University

May 2021

Professor Mary Billings
Professor Marti G. Subrahmanyam — Faculty Adviser

Professor Stanley Zin — Thesis Advisor

1. Introduction

The foreign exchange market is the largest financial market in the world, with an average daily volume of over $6 Trillion. Its participants range from tourists in Italy to the largest financial institutions in the world. Ask a foreign customer of a coffee shop in Florence how much a cup of coffee from New York would cost and why, he or she might look at you like you have four eyes. Given the implications for international economics, and finance, one would think that if the same question were posed to an economist, their eyes would light up. Surprisingly though, the economist might react similarly to the tourist – but for a different reason. As David Backus once said, “foreign exchange is where economic theory goes to die.”

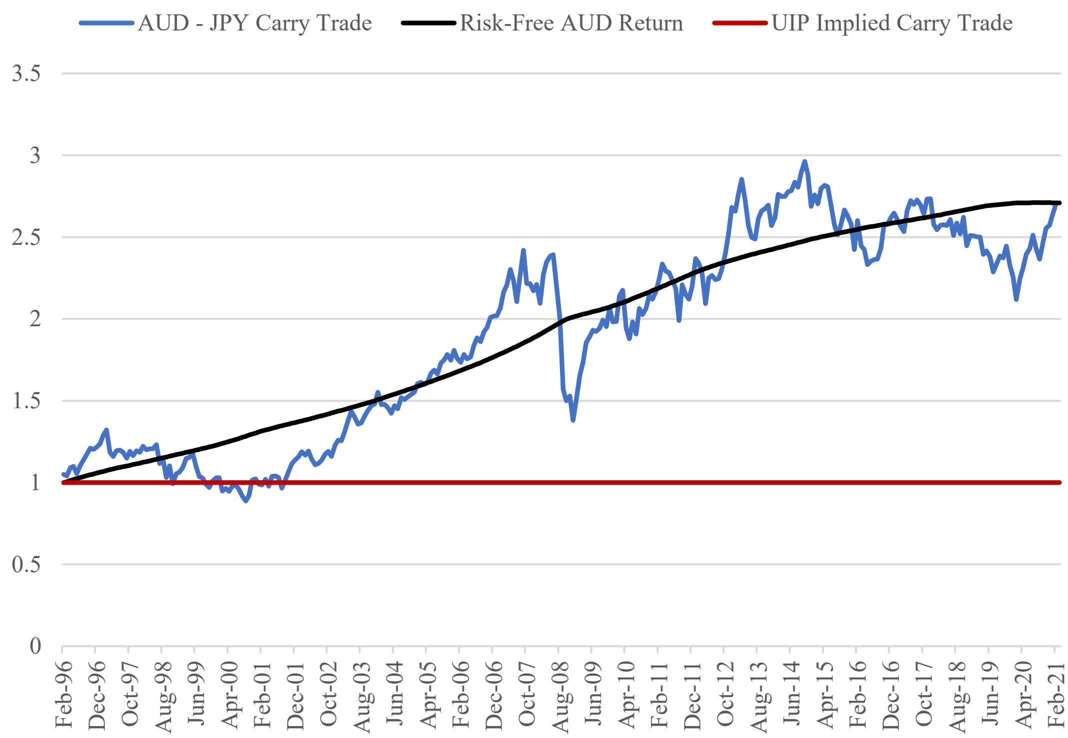
Economists have indeed spent many years trying to understand and explore the behavior of exchange rates in the context of economic theory. Although much progress has been made, equally many ‘puzzles’ and questions have been discovered. Nearly four decades ago, Meese and Rogoff (1983) famously argued that exchange rate behavior was governed by randomness. Today, many puzzles relating to exchange rates are still actively being explored in research.

The failure of uncovered interest rate parity (“UIP”) is one such puzzle that arises while studying the behavior of exchange rates in the context of international asset markets. UIP refers to the condition that exchange rates depreciate enough to offset interest rate differentials across countries, which is a prediction of almost all standard macroeconomic models. The carry trade is a trading strategy that capitalizes on the failure of UIP by borrowing in a currency with a low interest rate and investing in a currency with a high interest rate. If the high interest rate currency does not depreciate by an amount equal to the size of the differential, the carry trade is profitable. The failure of UIP can therefore be equally restated as the existence of positive returns on the carry trade.

The carry trade is an extremely popular trade, and notably profitable on the average. Take Japan and Australia for example. Japan has had significantly lower interest rates then Australia for the past two decades, yet the average Australian Dollar’s appreciation against the Japanese Yen is not nearly enough to offset that differential (see Figure 1).

Figure 1:

Cumulative returns on a 1-month carry trade



The positive expected return on the carry trade is evidence of either arbitrage opportunities or sizable risk premiums. Given that carry-trade funds routinely go bankrupt, the latter is of more interest. Moreover, given that the assumption of an arbitrage free environment is an integral part of the groundwork on which any economic theory is built. The question therefore arises: Is the observed behavior of interest rates and exchange rates across countries consistent with a time-varying risk premium in an arbitrage-free environment? If so, what is the nature of the risk that is driving such large risk premiums?

To discipline our answer to this question, we begin with the implications of an arbitrage free environment. Consider a dynamic environment where the world can enter various states $z \in$ Ƶ. If this environment is arbitrage free, then there exist Arrow securities – which have a payoff of one unit of consumption if a specific state z occurs and 0 if any other state occurs - for every state $z \in$ Ƶ. As such, the prices of these Arrow securities are often referred to as “state prices”.

Financial assets are defined by their payoffs in each state and, hence, any financial asset can be understood as portfolio of Arrow securities. In an arbitrage free environment, the price of any two assets with the same payoff structure must be equivalent. Thus, any asset in an arbitrage-free environment will have a price that is a combination of Arrow securities prices (i.e. the state prices).

Macroeconomic theory provides some direction in thinking about the determination of these state prices, often relating them to the marginal rate of substitution of a representative agent. Given preferences of a representative agent, the marginal rate of substitution represents the relative value of future consumption, in each future state, in terms of current consumption. It is the price, in today's consumption units, that the representative agent is willing to pay for one unit of future consumption in that state. This price depends crucially on the level of consumption endowed to the representative agent today and in the future. Because the level of consumption depends on the state of the world z, the marginal rate of substitution does as well. Therefore, if the level of consumption was known for every future state $z \in \mathbb{Z}$, one can map the marginal rate of substitution in each state to the state prices. In the simplest case, suppose the state of the world today and in the future was known exactly. Then the price of the Arrow security that pays off in the known state would be exactly equal to the marginal rate of substitution of the representative agent in that state of the world. In the case where the state of the world is unknown, the state prices would be equal to the marginal rate of substitution in that state but weighted by the probability of observing such a state. As such, the marginal rate of substitution is often called the pricing kernel, indicating the role it plays in 'pricing' assets.

Denote $Q_t^i(z)$ as the time t price of an Arrow security, which pays one unit of currency i as time $t+1$, if an only if state z occurs. $Q_t^i(z)$ is equal to the representative agent's nominal marginal rate of substitution $m_{t+1}^i(z)$, given state z occurs at time $t+1$, weighted by the probability $p(z)$ that state z occurs. The state prices are therefore:

$$Q_t^i(z) = m_{t+1}^i(z)p(z).$$

Consider the same notation for another currency j and denote $S_{t+1}(z)$ as the rate that currency j can be converted into one unit of currency i at time $t+1$, given state z occurs. $S_{t+1}(z)Q_t^j(z)$ would then be the price, in currency j, of receiving one unit of currency i time $t+1$, given state z occurs. In the absence of arbitrage, the currency i price of such a strategy would have to be equivalent to the price of the Arrow security which also pays one unit of currency i in state z:

$$\frac{1}{S_t} S_{t+1}(z)Q_t^j(z) = Q_t^i(z).$$

Thus, a complete markets assumption – that is, the existence of an arrow security for each state $z \in \mathbb{Z}$ – and a no-arbitrage assumption, yields a direct relationship between pricing kernels and the exchange rate:

$$\frac{S_{t+1}}{S_t} = \frac{m_{t+1}^i}{m_{t+1}^q}.$$

Backus, Foresi and Telmer (2001) ('BFT') detail further the general structure relating pricing kernels to exchange rates. They work with affine term structure models - specifically focusing on pricing kernels which incorporate normal (gaussian) risk. Much has changed in the two decades since that paper was written, especially the nature of risk. Financial crises, currency crashes, government collapses, and now natural disasters such as wildfire, hurricanes and pandemics are ubiquitous. Such rare but significant events suggest that risk is generally not normal and, the depth and breadth of these events suggest that non-normality arising from the arrival of disasters can't be overlooked. In the asset pricing literature, Barro (2009) and Reitz (1988) pioneered the incorporation of disaster risk alongside gaussian risk. Later, Wachter (2013) worked disaster risk that is persistent in nature.

In what follows, I consider a discrete-time affine term structure model like BFT. However, beyond normal risk, I incorporate persistent disaster risk and study its implications for exchange rates and interest rates across countries. The persistent nature of disaster risk is important. Rather than arising from the same probability distribution at each point in time, the distribution changes over time in a structured way. This reflects the environment of the economy. If the global economy is in a precarious position, it is unlikely to quickly leave that environment. Climate change is an excellent example. The same underlying issues that cause forest fires are likely the same ones that cause destructive hurricanes. Global financial crises are similar. Underlying fragilities of the global financial system can cause sequential crises, or a crisis that lasts many months, such as the great financial crisis and euro area crisis. There is an argument to be made, therefore, that there are extended periods where a disaster occurrence is more likely to occur and other periods where the opposite is true.

Unsurprisingly, global disasters have a great deal of influence on asset prices. The potential of global disasters increases risk worldwide and therefore can reduce risk free rates globally. In fact, one can argue that an increase in the potential of disaster has been one of the contributing factors to the trend of declining global interest rates. Equally, Wachter (2013) shows that persistent consumption disasters can explain many empirical features of equity returns like excess volatility and elevated equity premiums.

This paper applies similar reasoning to exchange rates. The notion of disasters in exchange rates is an age-old idea dating back to the origins of "the peso problem." Such a problem is inherently one-sided and affects only one country in a currency pair. Rather than focus on isolated rare events, I consider a global disaster, the likelihood of which is persistent and time-varying, that affects pricing kernels in asymmetric ways through heterogenous prices of disaster risk. This results in a pass through of global disasters to jumps in exchange rates. Countries with a relatively larger price of disaster risk have currencies that tend to appreciate sharply in a global disaster event, and relatively lower interest rates.

Naturally, the model generates a carry trade that may seem profitable. Borrowing in currencies with lower interest rates and investing in currencies with higher interest rates has crash risk like that outlined in Brunnermeier, Nagel and Pederson (2008). They model the crash risk as an unwinding of the carry trade, whereas the global disaster is an exogenous shock. Whether this disaster occurs in-sample or not, Uncovered Interest Rate Parity (UIP) does not hold. That is, there are positive expected returns on the carry trade.

In fact, a positive probability of a global disaster, and asymmetric prices of disaster risk are sufficient conditions to generate positive expected returns on the carry trade. Through the persistent nature of the probability of the disaster, these expected returns are both time-varying and persistent. Expected returns are largest (smallest) when the likelihood of disaster is high (low).

The basis of BFT was the forward premium anomaly, which refers to a negative correlation between the change in exchange rates and the interest rate spread – as was originally found in Fama (1984). UIP implies a correlation of 1. I find that while such a coefficient is consistently less than 1, it is rarely negative. BFT show that, to account for a coefficient less than zero – which I refer to as the "strong forward premium anomaly" – large asymmetries are required in the pricing kernels' loadings on the state variables. The same is true of the model presented here. However, such asymmetries are not necessary to generate a forward premium anomaly where the correlation is positive.

In the context of international asset markets, the model presented here illustrates the power of global disasters. A model that incorporates persistent time-varying global disaster risk, coupled with asymmetry in the price of that risk across pricing kernels, can generate: 1) positive expected returns on the carry trade along with its negative skewness, 2) safe-haven like properties of certain currencies, 3) excess volatility of exchange rates, and 4) persistent interest rate differentials across countries.

The paper continues as follows: Section 2 presents the model and its implications. Section 3 describes the data used to document the carry trade in practice and test properties of the model. Sections 4 and 5 provide parameter estimates of a simplified version of the model, illustrating the consequences of asymmetry in the price of disaster risk. Section 6 uses a more structural model to explore whether these asymmetries could reflect differences in monetary policy rules across countries. Finally, Section 7 concludes with a discussion of these findings in the context of other empirical research on exchange rates.

2. The Model

2.1 State Space

The state of the global economy evolves according to:

$$z_{1t+1} = (1-\varphi_1)\delta_1 + \varphi_1 z_{1t} + \sigma_1 \varepsilon_{1t+1} - d_{t+1},$$

where ε_{1t+1}, and d_{t+1} are random shocks and $\delta_1, \sigma_1 > 0$ and $|\varphi_1| < 1$ are parameters. ε_{1t+1} is assumed to be normally and identically distributed across time with $E\left[\varepsilon_{1i}\varepsilon_{1q}\right] = 0$ for $i \neq q$ (iid). Conditional on j_{t+1}, $d_{t+1} \sim N(\theta j_{t+1}, \rho^2 j_{t+1})$ – which represents the arrival of a disaster. Here, $j_{t+1} \in \{0, 1, 2, \ldots\}$, represents both the presence ($j_{t+1} > 0$) and size of a disaster. The disaster affects both the conditional mean and variance of z_{1t+1} through parameters θ and ρ respectively. For convenience, assume $\theta > 0$, which implies a disaster at time $t+1$ reduces the conditional mean of z_{1t+1}. If $\theta = \rho = 0$, i.e. no disaster occurs, z_{1t+1} follows a standard AR(1) process. As such, φ_1 would control the persistence ($Corr[z_{1t+1}, z_{1t}]$) of z_{1t+1}.

The presence and size of a disaster (j_{t+1}) is governed by a Poisson distribution:

$$p_t(j_{t+1}) = e^{-\omega z^*_{2t}}(\omega z^*_{2t})^{j_{t+1}}/j_{t+1}!,$$

where $\omega > 0$ and $p_t(j_{t+1})$ denotes the probability that the disaster takes on some value j_{t+1} for $j_{t+1} \in \{0, 1, 2, \ldots\}$, conditional on all information available at time t. ωz^*_{2t} is often referred to as the intensity parameter of the potential disaster j_{t+1}. It governs the probability of disaster at any given time t $\left(Prob[j_{t+1} \neq 0|z^*_{2t}] = 1 - e^{-\omega z^*_{2t}}\right)$ as well the expected magnitude of disaster $(E[j_{t+1}|z_{2t}] = \omega z^*_{2t})$. If ωz^*_{2t} were constant, the probability of disaster and the expected size of the disaster would be constant across time. Here, ωz^*_{2t} is time-dependent, which permits the probability of disaster and expected size of disaster to change across time. ωz^*_{2t} moves through time in a stochastic and persistent way through the state variable z_{2t+1}, which follows a standard AR(1):

$$z_{2t+1} = (1 - \varphi_2)\delta_2 + \varphi_2 z_{2t} + \sigma_2 \varepsilon_{2t+1},$$

where ε_{2t+1} is a random shock, which is also assumed to be iid normal. δ_2, $\sigma_2 > 0$ and $|\varphi_2| < 1$ are parameters. $z^*_{2t} = max\{0, z_{2t}\}$. If z_{2t} remains sufficiently far away from 0, then $z_{2t} \approx z^*_{2t}$ with high probability.

With such a structure for z_{2t+1}, the probability of disaster is not only time-dependent, but it is also persistent. That is, if the likelihood of a forthcoming disaster is high today, and $\varphi_2 > 0$, then it will *likely* continue to be high tomorrow. The same is true of the expected magnitude of a potential disaster at any given time t. If at time t the expected size of a potential disaster arriving at time $t + 1$ is relatively large, then come time $t + 1$, the expected size of a potential disaster at $t + 2$ will also *likely* be large. φ_2 governs this persistence in both the probability and expected magnitude of the potential disaster. Over the entire history, δ_2 governs the unconditional expected size of the disaster:

$$E[j_{t+1}] = E[E[j_{t+1}|z_{2t}]] = \omega\delta_2.$$

It is critically the likelihood of disaster that is persistent and not the disaster itself. A disaster at time t does not necessarily imply a disaster will again occur at time $t + 1$. At a more general level, one can think of this model reflecting the nature of economic disasters in practice. Often, the 'ingredients' of a disaster are present, and the economy is in a fragile state, in which a spark or tipping point might actually initiate a crisis. When z_{2t+1} is relatively large, the economy is in a fragile state and the distribution from which j_{t+1} is picked is more likely to yield a disaster. By the persistence of z_{2t+1}, when the state of economy is in a fragile state, it is likely to remain there. However, being in a fragile state does not mean a disaster necessary occurs, which is exactly what this state space implies.

2.2 Asset Pricing

In an arbitrage-free environment, the log pricing kernel for currency i cash flows is given by:

$$\log m^i_{t+1} = a^i + \gamma^i_1 z_{1t} + \gamma^i_2 z_{2t} + \lambda^i_1 \varepsilon_{1t+1} - \lambda^i_2 d_{t+1} + \lambda^i_3 \varepsilon_{2t+1}, \qquad (1)$$

where $a^i, \gamma_1^i, \gamma_2^i, \lambda_1^i, \lambda_2^i$ and λ_3^i are constant parameters. At this level of generality, the notion of a disaster is still somewhat abstract in the sense that it is merely a shock to the state variable z_{1t+1}, which affects the conditional distribution of the pricing kernel. To provide a level of concreteness, define the disaster as a period in which the conditional expectation of the pricing kernel jumps. To the extent that the pricing kernel is reflective of state prices, such reasoning is borne out of the idea that a 'disaster' state ought to have a high price. Payoffs are most valuable to a representative agent in a state where a disaster occurs. This restricts $\lambda_2^i < 0$, such that observing a $j_{t+1} \neq 0$ will likely increase the pricing kernel. Setting $\rho = 0$ guarantees that observing a $j_{t+1} \neq 0$ will increase the pricing kernel if $\lambda_2^i < 0$. However, because $j_{t+1} \neq 0$ also affects the variance of d_{t+1}, $j_{t+1} \neq 0$ only increases the likelihood of a pricing kernel jump if $\lambda_2^i < 0$ and $\rho \neq 0$. Following the affine term structure literature, λ_2^i is referred to as country i's price of disaster risk.

Conditional on all relevant information at time t, the cumulant generating function of country i's pricing kernel (assuming $z_{2t}^* = z_{2t}$) is

$$\log E_t\left[e^{s \log m_{t+1}^{\mathrm{i}}}\right] = k_t^i(s) = s\left(a^i + \gamma_1^i z_{1t} + +\gamma_2^i z_{2t}\right) + \frac{s^2\left(\left(\lambda_1^i\right)^2 + \left(\lambda_3^i\right)^2\right)}{2}$$
$$+ \omega z_{2t}\left(e^{\left(-s\lambda_2^i\theta + \left(\lambda_2^i\right)^2\rho^2 s^2/2\right)} - 1\right),$$

where $E_t[X]$ is shorthand notation for $E[X|z_{1t}, z_{2t}]$. The one-period interest rate is:

$$r_t^i = -\log E_t\left[m_{t+1}^{\mathrm{i}}\right] = -\, k_t^i(1)$$
$$r_t^i = -a^{i*} - \gamma_1^i z_{1t} - \gamma_2^i z_{2t} - \omega z_{2t}\left(e^{-\lambda_2^{i*}} - 1\right),$$

where,

$$a^{i*} = a^i + \frac{\left(\lambda_1^i\right)^2 + \left(\lambda_3^i\right)^2}{2}$$
$$\lambda_2^{i*} = \lambda_2^i\theta - \left(\lambda_2^i\right)^2\rho^2/2.$$

The short-term interest rate in country i is composed of a normal component (i.e. conditioning $E_t\left[m_{t+1}^{\mathrm{i}}\right]$ on $d_{t+1} = 0$) and a disaster driven component. Through the latter, one can immediately see a potential explanation for the global trend of declining interest rates over the past three decades. The short-term rate's disaster component $-\omega z_{2t}(e^{-\lambda_2^{i*}} - 1)$ is decreasing in the stochastic probability of disaster ωz_{2t}. All else equal, if the likelihood of a global disaster has increased over the past three decades, interest rates across all countries would be lower.

Consider a second country and currency q and denote the exchange rate (currency q cost of one unit of currency i) as S_t. Following BFT, the log exchange rate return from $t\ to\ t+1$ can be written as follows:

$$s_{t+1} - s_t = \log m^{\mathrm{i}}_{t+1} - \log m^{\mathrm{q}}_{t+1},$$

where $s_{t+1} = \log(S_{t+1})$. Substituting for the two pricing kernels yields the following stochastic process for the change in exchange rate ("FX Returs"):

$$\begin{aligned} s_{t+1} - s_t &= (a^i - a^q) + (\gamma_1^i - \gamma_1^q)z_{1t} + (\gamma_2^i - \gamma_2^q)z_{2t} \\ &+ (\lambda_1^i - \lambda_1^q)\varepsilon_{1t+1} - (\lambda_2^i - \lambda_2^q)d_{t+1} + (\lambda_3^i - \lambda_3^q)\varepsilon_{2t+1}. \end{aligned} \quad (2)$$

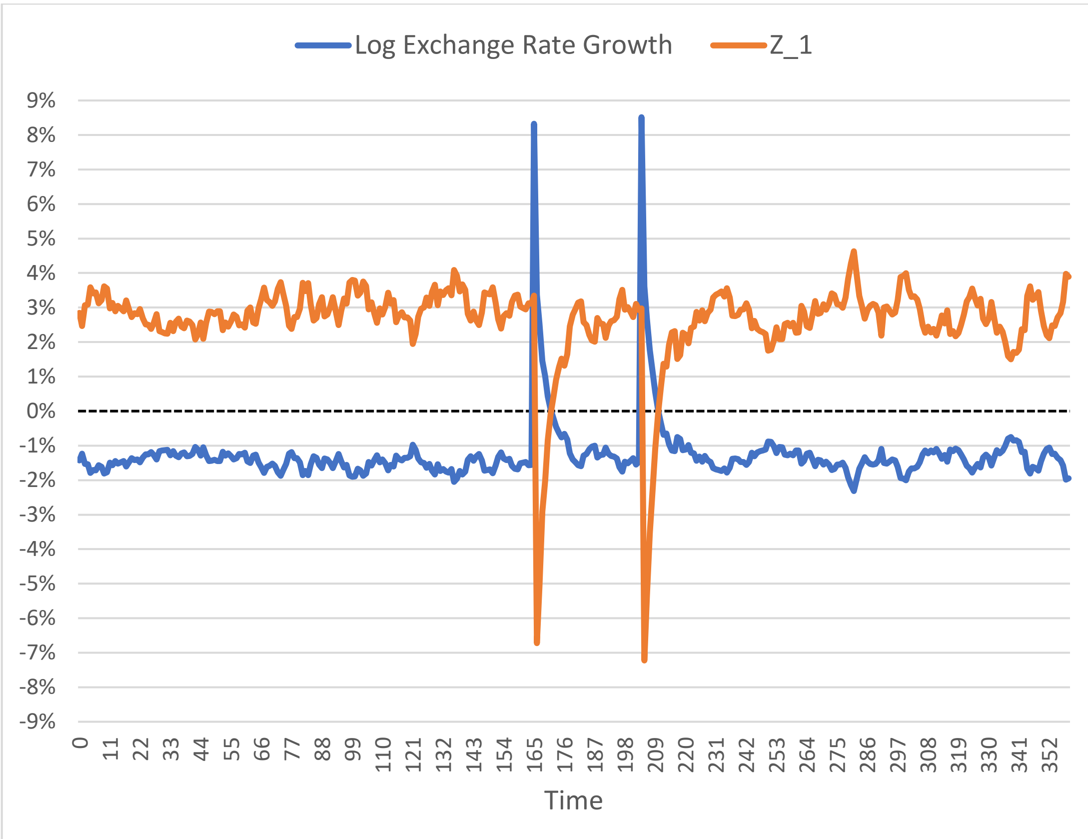
The change in the exchange rate process includes both normal and jump risk. The direction of the jump risk depends on the difference $\lambda_2^i\theta - \lambda_2^q\theta$. For simplicity, set $\rho = 0$, which implies the disaster does not affect the conditional variance of z_{1t}. This translates into substituting θj_{t+1} for d_{t+1} in
(2). In this case, if $\lambda_2^i < \lambda_2^q < 0$, then a disaster would lead to a sharp appreciation of currency i against currency q, all else equal.

To the extent that $\gamma_1^i - \gamma_1^q < 0$, (2) would also exhibit a mean reverting property like that described in Farhi and Gabaix (2016). To see this, assume z_{1t} is at its mean δ_1 and suppose a disaster occurs at time $t + 1$ ($j_{t+1} \neq 0$). All else equal, the exchange rate would jump higher and z_{1t+1} would jump below its mean. Now, suppose that for time $t + 2, t + 3 \ldots t + n$ moving forward that a disaster does not occur ($j_{t+2}, j_{t+3} \ldots j_{t+n} = 0$). Then, given $\gamma_1^i - \gamma_1^q < 0$, $s_{t+1} - s_t$ would fall as z_{1t+1} slowly increases and reverts towards its mean. To visualize this, consider the simulation in Figure 2.

Figure 2:

Reversion to Mean Simulation

Set $\gamma_1^i - \gamma_1^q = -.5$ and $\lambda_2^i\theta - \lambda_2^q\theta = -.1$. Take $\varphi_2 = .994, \delta_2 = .003, \sigma_2 = .00012$ and arbitrary values for the parameters governing z_{1t+1}, $\varphi_1 = .8, \delta_1 = .03, \sigma_2 = .003$ and $\theta = .1$. The chart below provides a simulation for z_{1t+1} and the FX process $s_{t+1} - s_t = -.5z_{1t} - .1j_{t+1}$. The high interest rate currency jumps when the disaster arrives. As the disaster subsides this appreciation decreases until the low interest rate currency appreciates after the disaster subsides.



Turning back to the interest rate, the cross-country interest rate differential is:

$$r_t^q - r_t^i = \log E_t[m_{t+1}^{\mathrm{i}}] - \log E_t[m_{t+1}^{\mathrm{q}}]$$

$$= a^{i*} - a^{q*} + (\gamma_1^i - \gamma_1^q)z_{1t} + (\gamma_2^i - \gamma_2^q)z_{2t} + \left(e^{-\lambda_2^{i*}} - e^{-\lambda_2^{q*}}\right)\omega z_{2t}. \quad (3)$$

Focusing strictly on disaster risk, (3) indicates that a larger price of disaster risk implies a lower short term interest rate across countries. If the two countires are symmetric in every way except the price of disaster risk, (i.e. $a^i = a^q, \gamma_1^i = \gamma_1^q,\ \gamma_2^i = \gamma_2^q, \lambda_1^i = \lambda_1^q$ & $\lambda_3^i = \lambda_3^q$) then

$$|\lambda_2^i| > |\lambda_2^q| \rightarrow r_t^q > r_t^i.$$

Thus, countries with a relatively larger price of disaster risk have 1) currencies that appreciate sharply when a disaster occurs *and* 2) relatively lower interest rates.

Gourio, Siemer and Verdelhan (2013) incorporate disasters into a two-country real business cycle model, but point to an important shortcoming. In the model, more 'risky' countries are those with lower interest rates. Their model suggests that these countries have more volatile quantities compared to high interest rate countries when the opposite is true in the data. Similarly, in this paper's model, low interest rate countries are riskier in the sense that they have a relatively higher price of disaster risk. However, the quantities and returns of low-rate countries are not necessarily more volatile. Once again, relax the assumption that $\gamma_1^i = \gamma_1^q$ and take the unconditional variance of the short rate for example:

$$Var[r_{1t}^i] = (\gamma_1^i)^2 Var[z_{1t}] + \left(\gamma_2^i + e^{\left(-\lambda_2^{i*}\right)} - 1\right)^2 Var[z_{2t}].$$

$Var[r_t^q]$ can be greater than $Var[r_t^i]$ - even if $|\lambda_2^i| > |\lambda_2^q|$ - given $(\gamma_1^q)^2 - (\gamma_1^i)^2$ sufficiently large.

2.3 The Carry Trade

Combining both the interest rate differential and the exchange rate process we can define a carry trade return - funding in currency i and lending in currency q - as follows:

$$c_{t+1} = r_{1t}^q - r_{1t}^i - (s_{t+1} - s_t).$$

What BFT refer to as the risk premium of currency q is essentially the conditional expected return on the carry trade:

$$E_t[c_{t+1}] = r_{1t}^q - r_{1t}^i - E_t(s_{t+1} - s_t)$$

$$E_t[c_{t+1}] = \left(\log E_t[m_{t+1}^{\mathrm{i}}] - E_t[\log m_{t+1}^{\mathrm{i}}]\right) - \left(\log E_t[m_{t+1}^{\mathrm{q}}] - E_t[\log m_{t+1}^{\mathrm{q}}]\right). \quad (4)$$

Therefore, the expected return on the carry trade is a function of the difference in conditional entropies of the two pricing kernels. Further, (4) is time varying and increasing in the conditional entropy of m_{t+1}^{i}. The conditional entropy of a random variable X is $\log E_t[X] - E_t[X]$. It follows from this definition, that the expected return on the carry trade can be written as the difference in higher order cumulants of the two pricing kernels:

$$E_t[c_{t+1}] = \sum_{j=2}^{\infty} \frac{k_{jt}^i - k_{jt}^q}{j!},$$

where k_{jt}^i denotes the $j'th$ cumulant of m_{t+1}^{i} conditional on z_{1t} and z_{2t}. In a strictly log-normal environment, $k_{jt}^i = 0$ for all $j > 2$. Thus, in such an environment the existence of positive returns on the carry trade restricts the second cumulant (i.e. the conditional variance) of m_{t+1}^{i} to be greater than that of m_{t+1}^{q}. Whereas in a model with disasters, the expected carry trade is a function of the (decaying) difference in variances as well as all other higher order conditional cumulants (such as those governing skewness and kurtosis). The infinite sum of such conditional cumulants can be written in closed form:

$$E_t[c_{t+1}] = \frac{\left(\lambda_1^i\right)^2 - \left(\lambda_1^q\right)^2 + \left(\lambda_3^i\right)^2 - \left(\lambda_3^q\right)^2}{2} + \omega z_{2t}\left(e^{\left[-\lambda_2^i\theta + \left(\lambda_2^i\right)^2\rho^2/2\right]} + \lambda_2^i\theta - e^{\left[-\lambda_2^q\theta + \left(\lambda_2^q\right)^2\rho^2/2\right]} - \lambda_2^q\theta\right).$$

The loadings on the state variables through γ_1^i and γ_2^i affect (2) and (3) equally and therefore drop out of the carry trade. That is, in a riskless world (or in a world with symmetric prices of risk) UIP would hold perfectly, as exchange rates and interest rates would react symmetrically to any change in the state of the world. It is risk, and asymmetric prices of that risk which generates the carry trade.

Once again, to isolate the effect of disaster risk, set $\lambda_1^i = \lambda_1^q$ and $\lambda_3^i = \lambda_3^q$, which implies

$$E_t[c_{t+1}] = \omega z_{2t}\left(e^{\left[-\lambda_2^i\theta + \left(\lambda_2^i\right)^2\rho^2/2\right]} + \lambda_2^i\theta - e^{\left[-\lambda_2^q\theta + \left(\lambda_2^q\right)^2\rho^2/2\right]} - \lambda_2^q\theta\right). \quad (5)$$

When $\rho = 0$, (5) simplifies to:

$$E_t[c_{t+1}] = \omega z_{2t}\left(e^{-\lambda_2^{i*}} + \lambda_2^{i*} - e^{-\lambda_2^{q*}} - \lambda_2^{q*}\right).$$

In both cases, asymmetry in the price of disaster risk across countries is sufficient to imply the existence of positive expected returns on the carry trade:

$$\left|\lambda_2^i\right| > \left|\lambda_2^q\right| \rightarrow E_t[c_{t+1}] > 0.$$

Therefore, any country pair (i, q) with $\left|\lambda_2^i\right| > \left|\lambda_2^q\right|$ would make for a strong candidate in a carry trade strategy. Note that (5) is increasing in ωz_{2t} for all such country pairs. Therefore, as the probability of a global disaster increases, the expected return on a carry strategy increases.

Critically, a global disaster need not occur for positive expected returns on the carry trade to exist. (5) is largest when the probability of disaster is highest, regardless of whether a disaster is observed or not. In fact, realized returns on the carry trade are strictly positive conditional on no disaster occurring. If $j_{t+1} = 0$, and prices of normal risk across countries are symmetric:

$$c_{t+1} = r_{1t}^q - r_{1t}^i - (s_{t+1} - s_t) = \omega z_{2t}\left(e^{-\lambda_2^{i*}} - e^{-\lambda_2^{q*}}\right),$$

which is greater than 0 when $\left|\lambda_2^i\right| > \left|\lambda_2^q\right|$. Given that the stochastic intensity parameter ωz_{2t} is time-varying, the realized return on the carry trade is as well, increasing in periods when a global disaster is more likely to occur (and crashing when it does occur).

Such a result fits nicely with the metaphor commonly attributed to the carry trade – that it is equivalent to "picking up pennies in front of a steamroller." (5) illustrates this feature explicitly. In fact, as one gets to the metaphorical steamroller (the greater the likelihood of seeing a disaster) one can expect to pick more pennies up (the higher the expected return on the carry trade), but with greater risk of being flattened by the steamroller (the greater the chance of observing a crash in the carry trade).

Often, the notion of a "safe haven currency" and the carry trade are linked. For example, the Japanese Yen and Swiss Franc are notoriously referred to as safe havens and are both popular funding currencies in the carry trade. In this environment, the model attributes an appropriate (and simple) mathematical interpretation to the term "safe-haven." That is, the pricing kernel of cash flows denominated in safe-haven currencies will have $|\lambda_2^i|$ relatively large and therefore disasters will increase the value of cash flows in that currency. As discussed earlier $|\lambda_2^i|$ relatively large also implies that disaster risk channels into lower rates. Thus, an environment with heterogenous prices of disaster risk can reproduce the real-world feature that safe haven currencies are those associated with low interest rate countries.

2.4 The Forward Premium Puzzle

The carry trade is also closely related to the forward premium puzzle, which is the focal point of BFT. Fama (1984) is perhaps the most influential study on the issues. The Fama regression refers to

$$s_{t+1} - s_t = b_0 + b_1\left(r_t^q - r_t^i\right) + residual,$$

where b_1 is often denoted the Fama coefficient. The forward premium puzzle refers to the empirical finding that b_1 is less than 1 and in fact (in studies at the time and prior to BFT) non-positive. $b_1 = 1$ and $b_0 = 0$ would imply that UIP holds and the expected return on the carry trade would be zero. As such, $b_1 < 1$ implies a failure of UIP while $b_1 < 0$ is a stronger statement. The model presented here can reproduce both, however, as will be shown in the data, b_1 is not consistently negative.[1]

As such, a simple but flexible affine model with asymmetry in the price of disaster risk is capable of capturing a wide range of international asset market behavior. To summarize, the heterogeneity in the price of disaster risk across pricing kernels implies a triple coincidence of 1) funding currency in the carry trade, 2) safe haven status and 3) positive exchange rate skewness. These features, amongst others, are worth exploring, and that is where the data comes in to play.

3. Data

Data for one month interest rate differentials and currency returns for four different country pairs is from Bloomberg. In each of these pairs, Japan or Switzerland serve as a low interest rate country while Australia and New Zealand represent high interest rate countries. The desire to capture funding and lending candidates in the carry trade motivates the use of these countries, which represent a wide range of interest rates, yet still have well measured data. Table 1 displays the moments for both interest rate differentials and currency returns for the period beginning January 1997 until March 2021.

[1] Appropriate choices for the $\gamma_1's$ and $\gamma_2's$ can generate the strong forward premium anomaly.

The data (plotted in Figure 3 and Figure 4) point to several important and well cited empirical facts:

1. **The existence of the carry trade:**

On average, FX returns do not wipe out interest rate differentials across countries. For example, New Zealand has had – on average – a one month interest rate which is .38% larger than that of Japan. Yet, the average one month return on the JPYNZD currency pair has been roughly 0%.

As is shown in Table 2, the Fama coefficient is positive – and greater than one - for three out of the four pairs. The fact that b_1 is greater than one is simply reflective of the fact that the sample is dominated by sharp jump of the Swiss Franc and Yen in 2008. In the case of CHFAUD, b_1 is less than 0, providing the only evidence of the strong forward premium anomaly – where the jump in 2008 does not dominate enough to bring b_1 above 0. The dominance of the exchange rate jump is illustrated in Figure 5 for the Yen.

One might interpret $b_1 > 1$ for several countries as UIP holding on the average through disasters. Rather than the subsequent one period exchange rate wiping out the previous periods interest rate differential, major crashes wipe out the interest rate differential accumulated over many periods. However, these findings must be considered carefully. As is shown in Lothian and Wu (2011), the Fama regression is not a consistent indicator of the breakdown in UIP. The coefficient is highly sample dependent.

The key result is that the relationship between the interest rate differential and exchange is a complicated one and can't be summarized by the simple relationship implied by UIP.

2. **Excess volatility of exchange rates:**

Across all four country pairs, the sample standard deviation of the currency return is 20 to 30 times larger than that of the interest rate differential.

3. **The persistence of interest rates:**

The 1st order autocorrelation of the interest rate differential is close to one across all country pairs. Even at further horizons, the interest rate differential remains relatively persistent, as can be seen from the one-year autocorrelations. On the other hand, currency returns exhibit very little persistence.

4. **The 'safe-haven' effect and negative skewness of the carry trade:**

The data point to consistent positive skewness in the Japanese Yen and Swiss Franc exchange rates against the currencies of high interest rate countries. This positive skewness translates to negative skewness (i.e. crash risk) in the carry trade – as is shown in Table 3.

Figure 3:

One Month Change in Log Exchange Rates

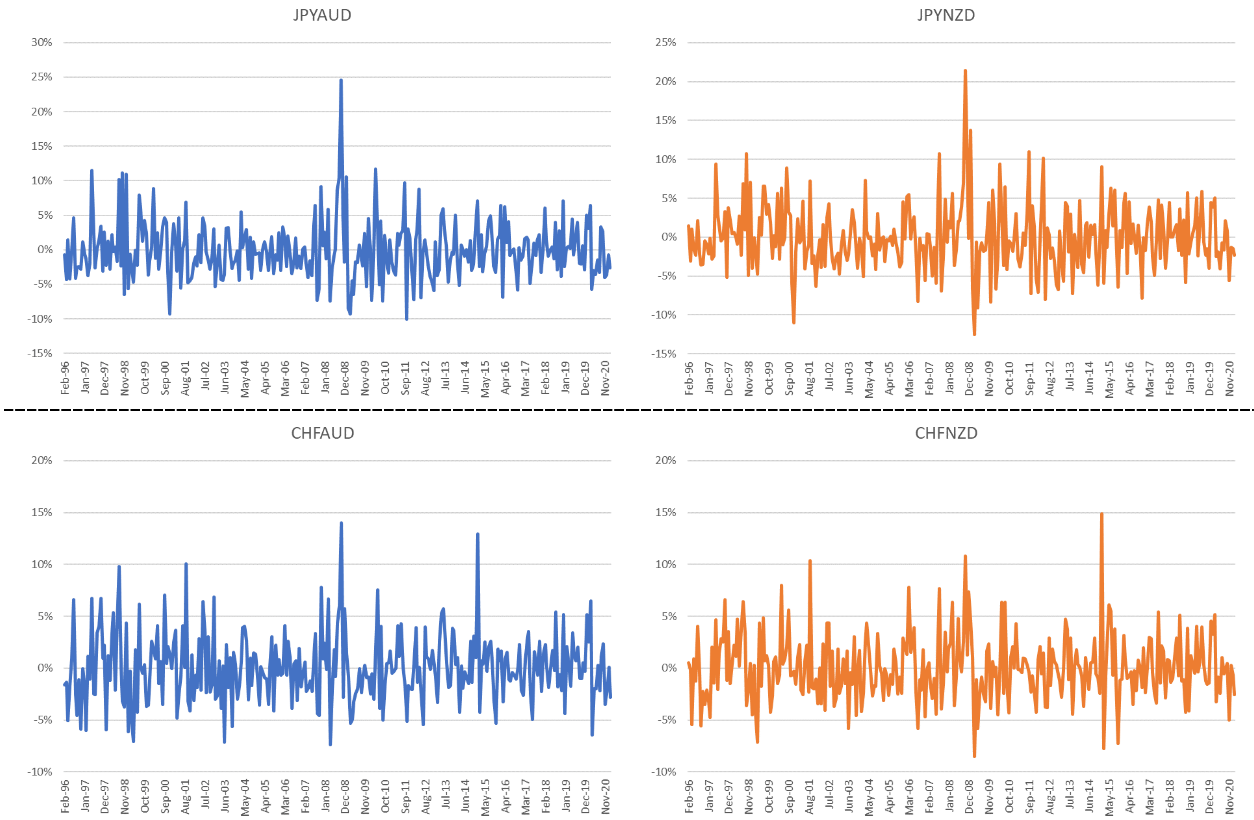
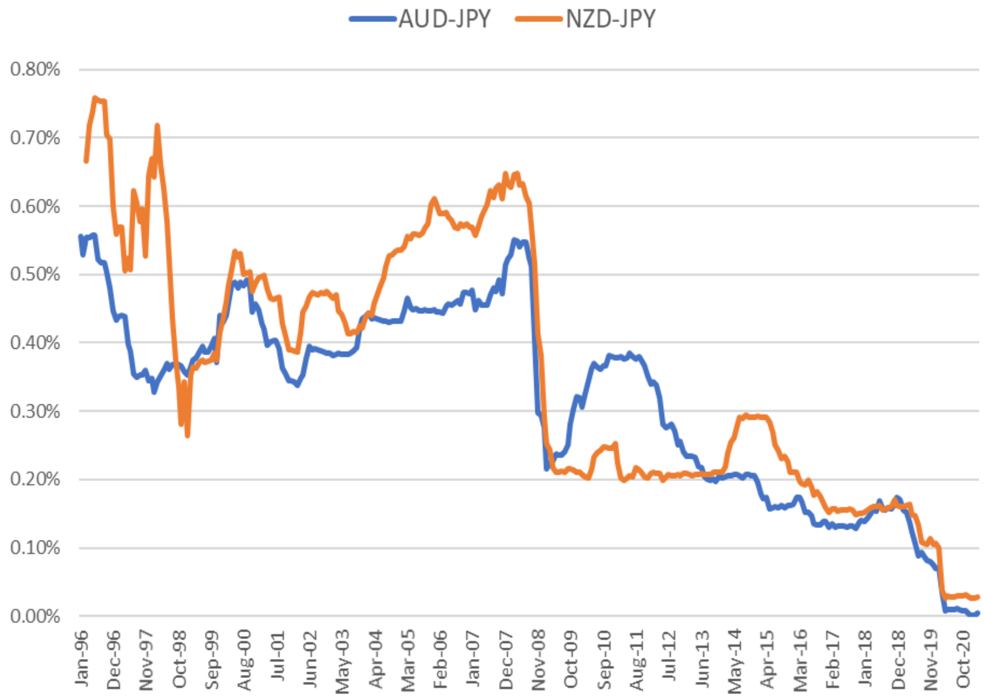


Figure 4:

One Month Interest Rate Differential



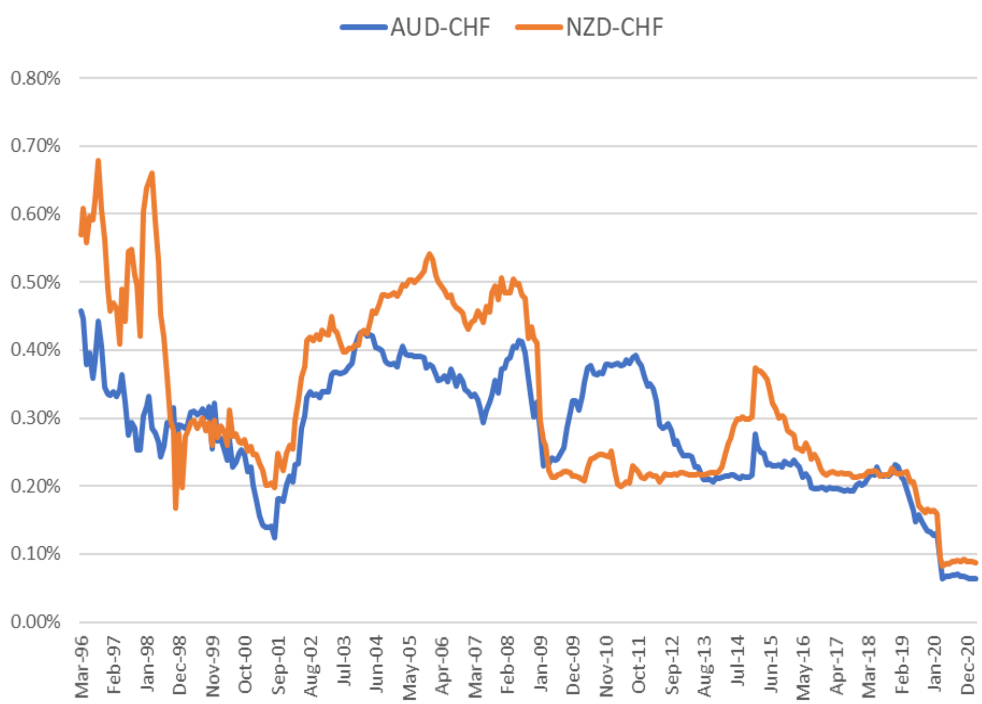


Table 1:[2]

Sample Moments: 1 Month FX Returns						
	Mean	***Standard Deviation***	***Skewness***	***Excess Kurtosis***	***1st Order Autocorrelation***	***12th Order Autocorrelation***
JPYAUD	*-0.017%*	*4.13%*	*1.07*	*3.89*	*0.008*	*0.005*
JPYNZD	*-0.024%*	*4.16%*	*.688*	*2.29*	*0.026*	*0.035*
CHFAUD	*0.077%*	*3.28%*	*.662*	*1.36*	*-0.022*	*-0.049*
CHFNZD	*0.069%*	*3.14%*	*.649*	*1.75*	*-0.011*	*-0.028*
Sample Moments: 1 Month Interest Rate Differentials						
AUD - JPY	*.325%*	*.148%*	*-.393*	*-.885*	*0.994*	*0.818*
NZD - JPY	*.369%*	*.198%*	*.208*	*-1.17*	*0.992*	*0.800*
AUD - CHF	*.289%*	*.094%*	*-.301*	*-.542*	*0.981*	*0.668*
NZD - CHF	*.332%*	*.142%*	*.510*	*-.744*	*0.980*	*0.696*

[2] Note: Skewness is calculated as $\hat{\mu}_3 / (\hat{\mu}_2)^{3/2}$ and Excess Kurtosis as $\frac{\hat{\mu}_4}{(\hat{\mu}_2)^2} - 3$. Where $\hat{\mu}_i$ is the i'th sample central moment.

Table 2:

Forward Premium Puzzle - Below is the estimated coefficient b_1 from the the Fama Regression: $s_{t+1} - s_t = b_0 + b_1(r_t^q - r_t^i) + residual$	
Country Pair	b_1
JPY // AUD	*1.63*
JPY // NZD	*1.37*
CHF // AUD	*-1.26*
CHF // NZD	*1.28*

Figure 5:

Fama Regression Scatter Plot

Plotted is the 1-month change in the log exchange rate against the prior months interest rate differential. The outliers from the great financial crisis are highlighted at the top in black.

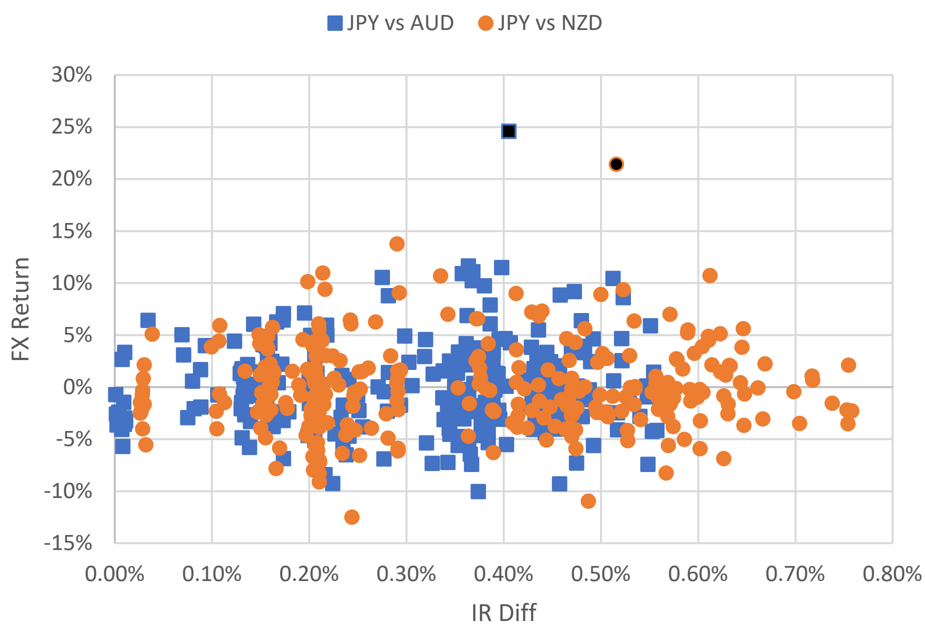


Table 3:

Skewness of the Carry Trade:	
Country Pair	***Skewness***
JPY // AUD	*-1.05*
JPY // NZD	*-.678*
CHF // AUD	*-.653*
CHF // NZD	*-.632*

4. Empirical Estimation

Estimating the parameters of the model for each of the four country pairs is a challenge. In its full form, the model has 21 parameters across two countries. The first step is to reduce the dimensionality of the model to a more manageable level.

First, set $\omega = 1$, which sets the scale for z_{2t}, and $\rho = 0$, which implies that the disaster does not change the conditional variance of z_{1t} and is strictly mean affecting through θj_{t+1}.

Next, assume symmetry in $a,\ \gamma_1,\ \lambda_1$ and λ_3 across each country pair. In words, equating these parameters implies that the effect of the of the state variable z_{1t} across the two pricing kernels is the same. That is, all behavior in the interest rate differential and exchange rate is driven by z_{2t}. Though this reduces the flexibility of the model in terms of the range of asset price behavior it can account for, doing so allows for explicit focus on the importance and effect of disaster risk.

These assumptions reduce the estimation problem to estimating 8 parameters, which is still a formidable task. In this version, the processes for the exchange rate and the interest rate differential simplify:

$$r_t^q - r_t^i = f_t - s_t = \left(\gamma_2^i - \gamma_2^q + e^{\left[-\lambda_2^i \theta\right]} - e^{\left[-\lambda_2^q \theta\right]}\right) z_{2t}$$

$$s_{t+1} - s_t = \log m_{t+1}^{\mathrm{i}} - \log m_{t+1}^{\mathrm{q}} = \left(\gamma_2^i - \gamma_2^q\right) z_{2t} - \left(\lambda_2^i \theta - \lambda_2^q \theta\right) j_{t+1}.$$

Further, rather than estimate δ_2, which governs the average probability of a disaster, δ_2 is fixed at four different values. Not only does doing so simplify the estimation, but it also provides a means to study how other parameters change as the probability of a global disaster varies. It also reflects the uncertainty surrounding the true probability of a rare global disaster. Wachter (2013) takes the probability of disaster to be 3.55% per year. This is in line with, but slightly more likely than the probability used by Gourio (2012) and in Barro et al. (2013), who estimate that a rare disaster occurs with a probability of 2.85% per year.

δ_2 is set to .001, .003, .005 and .008. $\delta_2 = .003$ implies that the average probability of a global disaster is roughly 3.55% per year, which is in line with Wachter (2013). To capture the possibility that a global disaster is rarer than what is implied by earlier estimates, δ_2 is also set to .001. This seems unlikely, however, as the sample used by Barro et al. (2013) and others to calculate the likelihood of disaster did not contain the current COVID-19 Pandemic. As such, δ_2 is additionally set to .005 and .008, implying a 5.82% and 9.15% average yearly probability of global disaster, respectively.

To further reduce the dimension, the difference in the $\gamma_2's$ across the two countries is estimated rather than the individual parameters themselves. The problem then simplifies to using GMM to estimate the following parameters: φ_2, σ_2, $\lambda_2^i\theta$, $\lambda_2^q\theta$ and $\gamma_2^i - \gamma_2^q$.

Given the goal of estimating the asymmetry of disaster prices across countries, it is worth noting that θ can not be separately identified from the $\lambda_2's$ in the model. The estimation therefore identifies $\lambda_2\theta$ for each country, from which an estimate for λ_2^i/λ_2^q can be derived. $\left(\lambda_2^i/\lambda_2^q\right) - 1$ reflects the relative difference in the price of disaster risk across countries.

In this version of the model, the change in the exchange rate process, and the interest rate differential are given by:

$$s_{t+1} - s_t = \left(\gamma_2^i - \gamma_2^q\right) z_{2t} - \left(\lambda_2^i - \lambda_2^q\right) d_{t+1}$$

$$r_t^q - r_t^i = \left(\gamma_2^i - \gamma_2^q\right) z_{2t} + \left(e^{-\lambda_2^{i*}} - e^{-\lambda_2^{q*}}\right) \omega z_{2t}.$$

The model is calibrated to match the following five moments: The mean, variance and 1st order autocorrelation of the interest rate differential as well as the mean and variance of the change in the exchange rate.

Denote the interest rate differential as $x_t = r_t^q - r_t^i$ and the log currency return as $y_t = s_{t+1} - s_t$. The mean, variance, and 1st order autocorrelation of the interest rate differential are given by (derivations in Appendix A):

$$E[x_t] = \left(\gamma_2^i - \gamma_2^q + e^{\left[-\lambda_2^i\theta\right]} - e^{\left[-\lambda_2^q\theta\right]}\right)\delta_2$$

$$Var[x_t] = \left[\left(\gamma_2^i - \gamma_2^q\right) + \left(e^{-\lambda_2^i\theta} - e^{-\lambda_2^q\theta}\right)\right]^2 \frac{{\sigma_2}^2}{1 - (\varphi_2)^2}$$

$$Corr[(x_t), (x_{t-1})] = \varphi_2.$$

The mean and variance of the log currency return are given by:

$$E[y_t] = \left(\gamma_2^i - \gamma_2^q - \lambda_2^i\theta + \lambda_2^q\theta\right)\delta_2$$

$$Var[y_t] = \left(\lambda_2^i\theta - \lambda_2^q\theta\right)^2\delta_2 + \left(\gamma_2^i - \gamma_2^q - \left(\lambda_2^i\theta - \lambda_2^q\theta\right)\right)^2 \frac{{\sigma_2}^2}{1 - (\varphi_2)^2}.$$

It is worth noting that: 1) The difference in the $\gamma_2{}'s$ affect the mean of the interest rate differential and log currency return in the same way; 2) The variance of the log currency return has an additional term from the conditional variance of the disaster; 3) The autocorrelation of the interest rate differential is governed entirely by the persistence of z_{2t}.

Using these moment restrictions, the exactly identified parameters and their standard errors[3] are reported for all 8 country pairs in Table 3. (See Appendix B for discussion of the GMM estimation and standard error calculations.)

[3] Standard errors are estimated using a Newey-West estimator with 12 lags.

Table 3A:

Australia									
JPYAUD					***CHFAUD***				
Parameter	Estimate				Parameter	Estimate			
δ_2	.001	.003	.005	.008	δ_2	.001	.003	.005	.008
φ_2	.989 (.023)	.989 (.059)	.989 (.137)	.989 (.254)	φ_2	.989 (.061)	.989 (.019)	.989 (.036)	.989 (.104)
σ_2	.00005 (.026)	.00015 (.022)	.00024 (.02)	.00039 (.018)	σ_2	.00006 (.024)	.00019 (.020)	.00032 (.017)	.00052 (.015)
$\lambda_2^{JPY}\theta$	-1.87 (.024)	-1.27 (.059)	-1.05 (.137)	-.877 (.254)	$\lambda_2^{CHF}\theta$	-1.58 (.063)	-1.06 (.021)	-.870 (.035)	-.720 (.103)
$\lambda_2^{AUD}\theta$	-.561 (.035)	-.519 (.025)	-.468 (.021)	-.415 (.019)	$\lambda_2^{AUD}\theta$	-.549 (.029)	-.464 (.021)	-.407 (.018)	-.354 (.016)
$\gamma_2^{JPY} - \gamma_2^{AUD}$	-1.48 (.053)	-.812 (.036)	-.619 (.030)	-.484 (.026)	$\gamma_2^{CHF} - \gamma_2^{AUD}$	-.268 (.044)	-.342 (.030)	-.310 (.025)	.-270 (.021)
$\frac{\lambda_2^{JPY}}{\lambda_2^{AUD}} - 1$	**2.30 (.213)**	**1.45 (.170)**	**1.24 (.316)**	**1.11 (.625)**	$\frac{\lambda_2^{CHF}}{\lambda_2^{AUD}} - 1$	**1.89 (.198)**	**1.29 (.121)**	**1.14 (.135)**	**1.04 (.307)**

Table 3B:

New Zealand									
JPYNZD					***CHFNZD***				
Parameter	Estimate				Parameter	Estimate			
δ_2	.001	.003	.005	.008	δ_2	.001	.003	.005	.008
φ_2	.992 (.008)	.992 (.030)	.992 (.066)	.992 (.120)	φ_2	.980 (.016)	.980 (.008)	.980 (.028)	.980 (.060)
σ_2	.00007 (.023)	.00020 (.019)	.00034 (.017)	.00054 (.016)	σ_2	.00008 (.028)	.00025 (.023)	.00042 (.020)	.00067 (.019)
$\lambda_2^{JPY}\theta$	-1.97 (.009)	-1.34 (.030)	-1.13 (.066)	-.945 (.120)	$\lambda_2^{CHF}\theta$	-1.75 (.017)	-1.20 (.007)	-.995 (.027)	-.830 (.060)
$\lambda_2^{NZD}\theta$	-.654 (.037)	-.598 (.023)	-.540 (.020)	-.479 (.017)	$\lambda_2^{NZD}\theta$	-.757 (.031)	-.627 (.024)	-.551 (.021)	-.480 (.019)
$\gamma_2^{JPY} - \gamma_2^{NZD}$	-1.56 (.049)	-.842 (.034)	-.638 (.029)	-.496 (.025)	$\gamma_2^{CHF} - \gamma_2^{NZD}$	-.298 (.065)	-.342 (.046)	-.305 (.039)	.-264 (.034)
$\frac{\lambda_2^{JPY}}{\lambda_2^{NZD}} - 1$	**2.02 (.171)**	**1.27 (.105)**	**1.09 (.148)**	**.972 (.264)**	$\frac{\lambda_2^{JPY}}{\lambda_2^{NZD}} - 1$	**1.31 (.102)**	**.913 (.075)**	**.805 (.081)**	**.731 (.133)**

5. Empirical Results

The simplified version of the model can successfully reproduce the first three of the four empirical facts mentioned earlier. The model generates positive expected returns on the carry trade, excess volatility of exchange rates and the persistence of interest rates across all currency pairs (and across all likelihoods of a global disaster).

5.1 Overestimating Skewness

Though the model captures the positive skewness of Yen and Franc returns (and thus the negative skewness of the carry trade), it significantly overstates the magnitude of that skewness.

The exchange rate process provides evidence for why this occurs. With $\delta_2 \leq .008$ the likelihood of seeing a $j_{t+1} > 1$ is essentially zero. Therefore, when a disaster occurs, currency i will appreciate against currency j by an amount equal to $\lambda_2^q\theta - \lambda_2^i\theta$ – which is much larger than any 1-month appreciation of the Yen or Franc within the sample. In fact, the largest one month move across all currencies took place in October 2008 when the Yen appreciated 25% against the Australian dollar. Whereas the model (in which $\delta_2 = .003$) would imply a JPYAUD 75% appreciation in the case of a disaster.

One can look at this as a shortcoming of the model. However, it is plausible (though somewhat sinister) that a true 'disaster' has yet to occur. Recall that with $\delta_2 = .003$, a disaster happens every 20 years with probability 50%. Under this assumption, the odds of seeing a disaster in the sample of data studied here is therefore 56% - which is far from a guarantee.

An additional (and more optimistic) explanation arises from the fact that in this simplified version, the price of normal risk (λ_1 and λ_3) is symmetric across countries. Thus, all the risk of the exchange rate process (2) is channeled through the asymmetry in the price of disaster risk λ_2. This asymmetry need not be as large if the assumption that $\lambda_1^i = \lambda_1^q$ and $\lambda_3^i = \lambda_3^q$ is relaxed and normal risk is allowed back in the model. Given that $\lambda_2^q\theta - \lambda_2^i\theta$ dictates the pass through of the disaster to the exchange rate, less asymmetry would lead to smaller jumps and less skewness in the exchange rate process.

Similarly, the assumption that disaster is strictly mean affecting $\rho = 0$ removes a source of risk (conditional variance) from the exchange rate. Relaxing this assumption would increase the power of disaster and reduce the burden that falls on the price of disaster risk.

If either or both of these latter two explanations are in fact true, the implication for the estimated model is the same: the asymmetry implied by this model likely overstates the true asymmetry.

5.2 Robustness

The model fails to yield precise estimates for the conditional variance $(\sigma_2)^2$ of the stochastic disaster intensity process z_{2t}. Whereas φ_2 – the parameter governing the persistence of z_{2t} and first order autocorrelation of the interest rate differential - is statistically significant in each estimation. The fact that the persistence of interest rate differentials is relatively similar across the various currency pairs points to the plausibility of a model with a single global state variable, rather than one with country specific state variables.

5.3 Forward Premium Anomaly

Though the parameters were not estimated to match the covariance of the exchange rate and interest rate differential or the Fama coefficient, estimates for both can be derived as:

$$Cov[x_t, y_t] = \frac{\left(\gamma_2^i - \gamma_2^q - \lambda_2^i\theta + \lambda_2^q\theta\right)\left(\gamma_2^i - \gamma_2^q + e^{\left[-\lambda_2^i\theta\right]} - e^{\left[-\lambda_2^q\theta\right]}\right)\frac{{\sigma_2}^2}{1-(\varphi_2)^2}}{\sqrt{Var[x_t]}\sqrt{Var[y_t]}}$$

$$b_1 = \frac{\gamma_2^i - \gamma_2^q - \lambda_2^i\theta + \lambda_2^q\theta}{\gamma_2^i - \gamma_2^q + e^{\left[-\lambda_2^i\theta\right]} - e^{\left[-\lambda_2^q\theta\right]}}. \qquad (6)$$

Table 4 shows implied and sample values for both. The simplified version of the model can capture both a positive and non-positive slope coefficient. However, the parameters identified seem to imply values of the correlation and Fama coefficient that are different from those observed in the data. In several cases, the sign is wrong, and in no case is the slope coefficient greater than one in the b_1 implied by the model.

Some intuition for why this is the case can be gained from (6). In the case where $|\lambda_2^i| > |\lambda_2^q|$, the only way to generate either $b_1 < 0$ or $b_1 > 1$ is for $\gamma_2^i - \gamma_2^q < 0$. That is,

$$|\lambda_2^i| > |\lambda_2^q| \; and \; \gamma_2^i - \gamma_2^q > 0 \; \rightarrow 0 < b_1 < 1.$$

If $e^{[-\lambda_2^q\theta]} - e^{\left[-\lambda_2^i\theta\right]} < \gamma_2^i - \gamma_2^q < \lambda_2^i\theta - \lambda_2^q\theta < 0$, then b_1 will be non-positive. In the case that $\lambda_2^i\theta - \lambda_2^q\theta < \gamma_2^i - \gamma_2^q < 0$, b_1 will be positive. However, if $\gamma_2^i - \gamma_2^q < e^{\left[-\lambda_2^i\theta\right]} - e^{[-\lambda_2^q\theta]} < 0$, then b_1 will be greater than one.

Therefore, to generate a Fama coefficient that is greater than one in this simplified version of the model, the difference in the loadings on stochastic disaster intensity parameter z_{2t} $\left(\gamma_2^i - \gamma_2^q\right)$ would have to be negative and sufficiently large – essentially overpowering the effect of the disaster on both the exchange rate and interest rate differential. To generate the strong forward premium anomaly $b_1 < 0$, the difference in the loadings on stochastic disaster intensity parameter z_{2t} would have to overpower the effect of the expected disaster on the exchange rate, but not the effect of the potential disaster on interest rates. This is evident in the implied means of the exchange rate and interest rate differential:

$$E[x_t] = \left(\gamma_2^i - \gamma_2^q + \left[e^{\left[-\lambda_2^i\theta\right]} - e^{[-\lambda_2^q\theta]}\right]\right)\delta_2$$

$$E[y_t] = \left(\gamma_2^i - \gamma_2^q + \left[\lambda_2^q\theta - \lambda_2^i\theta\right]\right)\delta_2.$$

$E[x_t] > 0$ and $E[y_t] < 0$ implies $b_1 < 0$ – which is precisely why the implied coefficients for JPYAUD and JPYNZD are negative. While $E[x_t] < 0$ and $E[y_t] < 0$ would imply that $b_1 > 1$. Given that for the sample covered, $E[x_t]$ is strictly positive, such a model will never generate $b_1 > 1$.

Table 4:

The Forward Premium Anomaly: Below are the model implied and observed values for the covariance of the exchange rate and interest rate differential as well as the Fama regression coefficient for all four country pairs in the case that $\delta_2 = .003$.

	Implied		Actual	
Country Pair	$Cov[x_t, y_t]$	b_1	$Corr[x_t, y_t]$	b_1
JPY // AUD	*-.002*	*-.052*	*.058*	*1.63*
JPY // NZD	*-.004*	*-.066*	*.065*	*1.37*
CHF // AUD	*.014*	*.267*	*-.036*	*-1.26*
CHF // NZD	*.017*	*.209*	*.057*	*1.28*

The difference in the loadings on z_{2t} clearly play a critical role in the relationship between the interest rate differential and the exchange rate – which is in line with the findings of BFT. In the estimation, $\gamma_2^i - \gamma_2^q$ is consistently negative and this finding is robust. $\gamma_2^i - \gamma_2^q < 0$ implies that as the probability of a rare global disaster increases, the interest rate differential widens – given $\left|e^{[-\lambda_2^i\theta]} - e^{[-\lambda_2^q\theta]}\right| > \left|\gamma_2^i - \gamma_2^q\right|$. It also implies the high interest rate currency appreciates as the risk of disaster increases. This is consistent with a time-varying risk premium explanation of the carry trade. As the risk of a global disaster increases, the expected return on the carry trade increases through both the interest rate differential and the exchange rate.

5.4 Asymmetry in Price of Disaster Risk

Perhaps the most striking and robust result of the estimation - and the main result of this paper - is the consistent finding that the low interest rate countries (Japan and Switzerland) have a larger price of disaster risk ($\lambda_2^i/\lambda_2^q > 0$) than the high interest rate countries (Australia and New Zealand). It is noteworthy that the size of the asymmetry needed to match the chosen moments decreases as the likelihood of the disaster (δ_2) increases, which is sensible. The more consequential the disaster is, the greater the impact of any level of asymmetry in the price of that disaster risk.

In this simplified version of the model, asymmetry in the price of disaster risk captures the triple coincidence of 1) funding currency in the carry trade, 2) safe haven status and 3) positive exchange rate skewness, all three of which are observed in the data through the Yen and Swiss Franc. Such a finding gives credence to the more general conjecture that disaster risk - and asymmetry in the price of such non-normal risk - can play a crucial role in understanding the observed behavior of exchange rates and interest rates in international asset markets.

6. A structural explanation

Until this point any economic structure behind the model, with the exception of the absence of arbitrage, has been abstracted away. However, to address the question of where the asymmetry in the price of disaster risk across countries might come from, it is necessary to add economic structure to the model.

Begin with a representative agent with Epstein and Zin (1989) recursive preferences:

$$U_t = \left[(1-\beta)c_t^{\eta} + \beta\mu_t(U_{t+1})^{\eta}\right]^{1/\eta}$$
$$\mu_t(U_{t+1}) = E_t[{U_{t+1}}^{\alpha}],$$

where $\alpha < 1$, $\eta < 1$ and $0 < \beta < 1$, are preference parameters capturing risk aversion, intertemporal substitution, and time-preference respectively. Assume the same stochastic processes as before and explicitly define z_{1t+1} as consumption growth:

$$z_{1t+1} = log\left(\frac{c_{t+1}}{c_t}\right)$$

In this environment, the real pricing kernel can then be written in reduced form:

$$\log m_{t+1}^{(R)i} = a^i + \gamma_1^i z_{1t} + \gamma_2^i z_{2t} + \lambda_1^i \varepsilon_{1t+1} - \lambda_2^i d_{t+1} + \lambda_3^i \varepsilon_{2t+1}$$

Where $a^i, \gamma_1^i, \gamma_2^i, \lambda_1^i, \lambda_2^i$ and λ_3^i are each nonlinear functions of the structural parameters of preferences along with parameters governing the dynamics of consumption growth.

Now consider a Taylor rule followed by country i's central bank of the following form:

$$r_t^i = \hat{\tau}^i + \tau_\pi^i \pi_t^i + \tau_z^i z_{1t} + \tau_d^i z_{2t}, \tag{6}$$

where π_t^i denotes the rate of inflation. As is standard, the Taylor rule depends on inflation π_t and output z_{1t}, through parameters τ_π and τ_z respectively. However, (6) also contains a policy discretion term $\tau_d^i z_{2t}$ related to the intensity parameter of a potential disaster. As such, τ_d^i governs the sensitivity of monetary policy at time t to the likelihood of a disaster arriving the following period.

To the extent that central banks react proactively to fragilities in the global economy, this is a fair representation of a policy rule. A central bank that would prefer to get "ahead of the curve" and adjust interest rates knowing there is a chance of a large shock to output growth might have $\tau_d^i \neq 0$.

In equilibrium, inflation adjusts so that the monetary policy rule is aligned with the bond market:

$$r_t^i = \hat{\tau}^i + \tau_\pi^i \pi_t^i + \tau_z^i z_{1t} + \tau_d^i z_{2t} = r_{1t}^i = -\log E_t \left[e^{\log m_{t+1}^{(R)i} - \pi_{t+1}^i} \right].$$

Using a guess and verify approach to solve for the equilibrium stochastic process for inflation yields:[4]

$$\pi_t = b_0^i + b_1^i z_{1t} + b_2^i z_{2t},$$

where:

$$b_0^i = -\frac{\left(\hat{\tau}^i + a^i + \left(\lambda_1^i - b_1^i \sigma_1\right)^2 + \left(b_2^i \sigma_2\right)^2 - b_1^i (1-\varphi_1)\delta_1 - b_2^i (1-\varphi_2)\delta_2\right)}{(\tau_\pi^i - 1)}$$

$$b_1^i = -\frac{\left(\tau_z^i + \gamma_1^i\right)}{(\tau_\pi^i - \varphi_1)}$$

$$b_2^i = -\frac{\left(\tau_d^i + \gamma_2^i + \left(e^{\left(\lambda_2^i - b_1\right)\theta} - 1\right)\right)}{(\tau_\pi^i - \varphi_2)}.$$

Thus, the equilibrium process for the inflation rate depends on deeper parameters of the model, including: 1) structural parameters of preferences; 2) parameters of the Taylor rule and 3) parameters governing the dynamics of consumption growth.

In the simplest case, preferences and therefore the parameters of the real pricing kernel are symmetric across countries. In other words, PPP holds, which implies:

$$s_{t+1} - s_t = \log m_{t+1}^{(R)i} - \pi_{t+1}^i - \left(\log m_{t+1}^{(R)q} - \pi_{t+1}^q\right) = \pi_{t+1}^q - \pi_{t+1}^i$$

$$s_{t+1} - s_t = \left(b_0^q - b_0^i\right) + \left(b_1^q - b_1^i\right)(z_{1t+1}) + \left(b_2^q - b_2^i\right)(z_{2t+1})$$

$$s_{t+1} - s_t = \left(b_0^i - b_0^q\right) + \left(b_1^q - b_1^i\right)\left((1-\varphi_1)\delta_1 + \varphi_1 z_{1t} + \sigma_1 \varepsilon_{1t+1} - d_{t+1}\right) + \left(b_2^q - b_2^i\right)\left((1-\varphi_2)\delta_2 + \varphi_2 z_{2t} + \sigma_2 \varepsilon_{2t+1}\right).$$

Now, to replicate a version of the model like the one utilized in the estimation, set $\rho = 0$. Additionally, impose $\varphi_1 = 0$. In words, this implies that consumption growth is iid, conditional on no disaster. In this case, the log currency return is:

$$s_{t+1} - s_t = \left(b_0^q - b_0^i\right) + \left(b_1^q - b_1^i\right)\delta_1 + \left(b_1^q - b_1^i\right)\sigma_1 \varepsilon_{1t+1} - \left(b_1^q - b_1^i\right)\theta j_{t+1} + \left(b_2^q - b_2^i\right)(1-\varphi_2)\delta_2 + \left(b_2^q - b_2^i\right)\varphi_2 z_{2t} + \left(b_2^q - b_2^i\right)\sigma_2 \varepsilon_{2t+1}.$$

This is precisely of the form:

[4] For derivations and more on this model see Gallmeyer et al. (2007).

$$s_{t+1} - s_t = \left(a^{\$i} - a^{\$q}\right) + \left(\gamma_2^{\$i} - \gamma_2^{\$q}\right) z_{2t}$$
$$+ \left(\lambda_1^{\$i} - \lambda_1^{\$q}\right) \varepsilon_{1t+1} - \left(\lambda_2^{\$i} - \lambda_2^{\$q}\right) \theta j_{t+1} + \left(\lambda_3^{\$i} - \lambda_3^{\$q}\right) \varepsilon_{2t+1}.$$

Here a superscripted \$ denotes that the parameter is of the nominal pricing kernel $\left(\log m_{t+1}^{\$i} = \log m_{t+1}^{(R)i} - \pi_{t+1}^{i}\right)$ – which is what has been used throughout this paper.

As such, an endogenous inflation model provides some intuition for where the asymmetry in the price of disaster risk comes from. In the case that $\varphi_1 = 0$, $\gamma_1^i = 0$ and $b_1^i = -\frac{(\tau_z^i)}{(\tau_\pi^i)}$. Thus, lining up coefficients implies:

$$\left(\lambda_2^{\$i} - \lambda_2^{\$q}\right) \theta = \left(b_1^q - b_1^i\right)\theta = \left(\frac{\tau_z^i}{\tau_\pi^i} - \frac{\tau_z^q}{\tau_\pi^q}\right)\theta$$
$$\left(\lambda_2^{\$i} - \lambda_2^{\$q}\right) = \left(\frac{\tau_z^i}{\tau_\pi^i} - \frac{\tau_z^q}{\tau_\pi^q}\right).$$

The existence of the carry trade (arising from asymmetry in the price of disaster risk $\lambda_2^{\$i} < \lambda_2^{\$q} < 0$) reflects asymmetry in the Taylor rules governing monetary policy across countries:

$$\lambda_2^{\$i} < \lambda_2^{\$q} < 0 \rightarrow \left|\lambda_2^{\$i}\right| > \left|\lambda_2^{\$q}\right| \rightarrow \frac{\tau_z^i}{\tau_\pi^i} < \frac{\tau_z^q}{\tau_\pi^q}$$

Clearly, monetary policy across both countries must respond to output (i.e. $\tau_z^i, \tau_z^q \neq 0$) for any asymmetry in the price of disaster risk to exist. Further, a country with an output-focused Taylor rule (relative to inflation) will have a lower observed price of disaster risk than that of a country whose monetary policy rule is more inflation focused (relative to output).

In the case of Japan and Australia for example, such a result yields three interpretations for the asymmetry in the price of disaster risk $\left(\frac{\lambda_2^{JPY}}{\lambda_2^{AUD}} - 1\right)$:

1) In the case that $\tau_z^{JPY} = \tau_z^{AUD}$, then the carry trade requires $\tau_\pi^{JPY} > \tau_\pi^{NOK}$. That is, the Bank of Japan must be a more aggressive inflation targeting central bank then the Reserve Bank of Australia.
2) In the case that $\tau_\pi^{JPY} = \tau_\pi^{AUD}$, then the carry trade requires $\tau_z^{JPY} < \tau_z^{AUD}$. That is, the Reserve Bank of Australia responds more to global output in setting monetary policy then does the Bank of Japan.
3) A combination of 1 and 2.

These results generalize to the other 3 pairs estimated earlier. The similarity of Japan and Switzerland as 'low interest rate countries' in this model, is interestingly the result of following similar policy rules. It must be, relative to other high interest rate countries that Switzerland and Japan, share one of the three results presented above. Given that these two developed countries were two of the first to test the waters of the so-called Zero Lower Bound, such a result seems plausible.

Though the asymmetry of the price of disaster risk is the focal point of this paper, it is worth noting that each estimate of $\gamma_2^{\$i} - \gamma_2^{\$q}$, where i represents Japan or Switzerland or q represents Australia or New Zealand, was found to be negative. As was the case for the asymmetry in price of disaster risk, different sensitivities to output and inflation across central bank policy rules can equally drive asymmetry in the $\gamma_2^{\$}$'s. However, the endogenous inflation model illustrates that even if central banks responded to output and inflation symmetrically across countries, asymmetry in the sensitivity of policy rules to the time-varying probability of disaster can drive a wedge between $\gamma_2^{\$i}$ and $\gamma_2^{\$q}$.

To illustrate this, the coefficient matching process used above is employed once again:

$$\gamma_2^{\$i} - \gamma_2^{\$q} = \left(b_2^q - b_2^i\right)\varphi_2.$$

If central banks across countries are equally sensitive to output and inflation (i.e. $\tau_\pi^i = \tau_\pi^q$ and $\tau_z^i = \tau_z^q$), then $b_2^q - b_2^i$ reduces to:

$$b_2^q - b_2^i = \frac{\tau_d^i - \tau_d^q}{(\tau_\pi^i - \varphi_2)}. \tag{7}$$

Given realistic Taylor rule parameters, that is $\tau_\pi^i = \tau_\pi^q > 1$, and $|\varphi_2| < 1$, the denominator of (7) will be strictly positive and

$$\tau_d^i - \tau_d^q < 0 \rightarrow b_2^q - b_2^i < 0 \rightarrow \gamma_2^{\$i} - \gamma_2^{\$q} < 0.$$

That is, asymmetry in $\gamma_2^{\$}$ across countries would be driven solely by asymmetry in the sensitivity across central bank policy rules to the likelihood of a disaster.

If central banks did indeed respond to the probability of a global disaster, they would likely lower (raise) rates to get ahead of the curve in the case that the probability of disaster increased (decreased). A reasonable assumption, therefore, is that $d < 0$. As such, in the case that monetary authorities respond symmetrically to changes in output and inflation, $\gamma_2^{\$i} - \gamma_2^{\$q} < 0$ reflects the fact that $|\tau_d^i| > |\tau_d^q|$. Thus, one possible explanation for $\gamma_2^{\$i} - \gamma_2^{\$q}$ being consistently negative in estimation is that the Bank of Japan and Swiss National Bank are perhaps more sensitive to the time-varying probability of a forthcoming global disaster in setting monetary policy, then are the Reserve Bank of Australia and Reserve Bank of New Zealand.

7. Related Work

The basis of this paper is Backus Foresi and Telmer (2001). There the set up relating nominal pricing kernels and exchange rates is presented, and a class of affine term-structure models with normal risk are evaluated. Much of their work is devoted to assessing the strong premium forward anomaly, whereas in the sample of data evaluated, UIP breaks down but not to the extent that the Fama regression coefficient is negative in all cases – as Lothian and Wu (2011) also document.

This paper is also related to the long strand of literature looking at disasters and asset prices. Starting from Reitz (1988) and Barro (2008) up until Wachter (2013), whose continuous time persistent disaster risk model is discretized here. Gourio (2012) looks at disasters in the context or a real business cycle model, illustrating the importance disaster risk plays in business cycles and asset prices. Backus, Chernov and Martin (2011) look at option prices to quantify the distribution of disaster risk priced in Equity options. Nakamura et al. (2013) attempt to estimate the likelihood of observing a consumption disaster by looking at a sample of many countries and may periods. These three papers serve as the basis for choosing δ_2 in the simplified version of the model estimated here. All conjecture that the probability of a rare disaster event occurring is less than 5% per year.

In looking specifically at exchange rates and the carry trade, this paper falls into the subset of asset pricing literature focusing on exchange rates. Lustig, Roussanov, and Verdelhan (2011) studied the carry trade explicitly and found that heterogeneity in exposure to a global factor was sufficient in generating positive expected excess returns on the carry trade. This paper similarly focuses on a global factor, but – here - it is critically asymmetry in the price of global disaster risk (not normal risk) that generates positive excess returns on the carry trade. Brandt, Cochrane, Santa-Clara (2006) demonstrate that international risk sharing takes place and pricing kernels across country comove. The global structure of the sate variables within the model presented here reflect such a finding.

Lastly, this paper relates to the literature that incorporates rare disasters in trying to understand the behavior of exchange rates. The well-developed "Peso-Problem" literature dates to Kaminsky (1993) and Engel and Hamilton (1990). The idea of compensation for a 'one off' event is not a new one. However, the Peso Problem is a generally thought of as a one-sided problem – capturing a large idiosyncratic shock. This paper specifically looks at global disasters – affecting all countries.

More recently, Chernov, Graveline and Zviadadze (2018) study empirically the importance of crash risk in currency returns. Though they look at idiosyncratic jump risk, the model presented here is restricted to global disaster risk. Brunnermeier, Nagel, and Pedersen (2008), propose that it is the unwinding of the carry trade itself which creates crash risk in currencies. They show that currencies that are strong candidates as an investment currency in the carry trade are most exposed to such crash risk. To the extent that an 'unwinding' takes place in periods I which a global disaster occurs, the model presented here can capture their findings.

Gourio, Siemer and Verdelhan (2013) present a real framework and illustrate that heterogenous exposure to a global disaster through a productivity shock can generate the failure of UIP. Their model implies that low interest rate countries are riskier, in the sense that they are more exposed to the global disaster. Horvath (2020) provides merit to such a result, showing that it may be true that advanced economies with lower rates are more exposed to global economic disasters. However, their model also implies that low interest rate countries have more volatile quantities, which seems to disagree with the data. Such an issue does not necessarily constrain the model presented here, even though low rate countries are more 'exposed' to the disaster in that they have a higher price of disaster risk.

Farhi and Gabaix (2016) also apply rare disasters in the context of exchange rates, and they too endogenize the disaster through the productivity of the export sector. They find similar results to this paper, in that incorporating global disasters into a model provide the power to account for a large portion of the observed behavior of exchange rates.

In this paper, initially, the channel through which the disaster generates positive excess returns of the carry trade occurs is unspecified. The pass through from the global disaster to the exchange rate is simply the result of asymmetry in the price of disaster risk in the nominal pricing kernels across countries. The cause of this asymmetry is, at first, abstracted away from. Both Gourio, Siemer and Verdelhan (2013) and Farhi and Gabaix (2016) illustrate one such possible cause. That is, heterogenous exposure to the global disaster through a productivity shock. This paper presents another possible source of the asymmetry in the price of disaster risk through asymmetry in the parameters governing the monetary policy rules across countries.

8. Conclusion

Excess returns on the carry trade likely reflect sizeable and time varying risk premiums. In this paper, the magnitude of the risk premium reflects compensation for non-normal risk, which arises from a global disaster. The return varies through time with the time-varying probability of a global disaster. The jump in the exchange rate that triggers the crash in the carry trade arises from the fact the price of global disaster risk is asymmetric across currencies. Currencies with a relatively higher price of disaster risk appreciate sharply (payoffs in that currency are preferred) when a disaster occurs and have lower interest rates – which is a natural representation of the carry trade in practice.

Though one explanation is presented here, there are other plausible sources for the asymmetry in the price of disaster risk across countries. Here, the asymmetry arises from different monetary policy rules across countries. Monetary authorities need not respond to the changing probability of a global disaster for such a result to hold. However, the setup of the model provides flexibility to think about how the monetary authority might respond to the likelihood of disaster in setting interest rates. Exploring the reasonability of such a rule and its implications for quantities and asset prices is an interesting direction of future research, but it is well outside the scope of this paper.

Generally, the work presented here points to the potential importance of disasters in understanding the behavior of asset prices. Work on this subject is by no means new, but it has been revitalized in recent years. Specifically, within the application to exchange rates, there are two questions that may deserve further research. 1) Identifying and distinguishing the effect of country specific (idiosyncratic) disasters and global disasters. 2) Identifying and distinguishing the contribution of disaster risk relative to normal risk. In the model presented here, prices of normal risk were assumed to be symmetric across countries for simplicity, but that is unlikely a realistic assumption.

Another interesting area for future research would explore how a time-varying probability of a global disaster affects real quantities. Understanding the nature of global disasters and the channels through which they alter real variables is critical. Further, to the extent that global disasters are consequential, understanding what causes global disasters is of equal importance.

Lastly, future work could identify and estimate structural models of preferences and policy. If asset prices reflect preferences and policy, then the behavior of asset prices yields insight into their structure. Considering this, the work presented here points to a potential application of exploring whether asset behavior could reflect risk preferences that are even more sensitive to disasters, such as disappointment aversion.

References

Arrow, Kenneth J., and Gerard Debreu. "Existence of an Equilibrium for a Competitive Economy." *Econometrica* 22, no. 3 (1954): 265. doi:10.2307/1907353.

Backus, David K., Silverio Foresi, and Chris I. Telmer. "Affine Term Structure Models and the Forward Premium Anomaly." *The Journal of Finance* 56, no. 1 (2001): 279-304. doi:10.1111/0022-1082.00325.

Backus, David, Mikhail Chernov, and Ian Martin. "Disasters Implied by Equity Index Options." *The Journal of Finance* 66, no. 6 (2011): 1969-2012. doi:10.1111/j.1540-6261.2011.01697.x.

Barro, Robert J. "Rare Disasters, Asset Prices, and Welfare Costs." *American Economic Review* 99, no. 1 (2009): 243-64. doi:10.1257/aer.99.1.243.

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* 53, no. 4 (2006): 671-98. doi:10.1016/j.jmoneco.2005.02.004.

Brunnermeier, Markus K., Stefan Nagel, and Lasse H. Pedersen. "Carry Trades and Currency Crashes." *NBER Macroeconomics Annual* 23, no. 1 (2008): 313-48. doi:10.1086/593088.

Chernov, Mikhail, Jeremy Graveline, and Irina Zviadadze. "Crash Risk in Currency Returns." *Journal of Financial and Quantitative Analysis* 53, no. 1 (2018): 137-70. doi:10.1017/s0022109017000801.

Engel, Charles, and James Hamilton. "Long Swings in the Exchange Rate: Are They in the Data and Do Markets Know It?" *American Economic Review* 80(4) (1990): 689-713. doi:10.3386/w3165.

Epstein, Larry G., and Stanley E. Zin. "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework." *Econometrica* 57, no. 4 (1989): 937. doi:10.2307/1913778.

Fama, Eugene F. "Forward and Spot Exchange Rates." *Journal of Monetary Economics* 14, no. 3 (1984): 319-38. doi:10.1016/0304-3932(84)90046-1.

Farhi, Emmanuel, and Xavier Gabaix. "Rare Disasters and Exchange Rates *." *The Quarterly Journal of Economics* 131, no. 1 (2015): 1-52. doi:10.1093/qje/qjv040.

Gallmeyer, Michael F., Burton Hollifield, Francisco J. Palomino, and Stanley E. Zin. "Arbitrage-Free Bond Pricing with Dynamic Macroeconomic Models." *Federal Reserve Bank of St. Louis Review*, August 2007, 305-26. doi:10.20955/r.89.305-326.

Gourio, François. "Disaster Risk and Business Cycles." *American Economic Review* 102, no. 6 (2012): 2734-766. doi:10.1257/aer.102.6.2734.

Gourio, François, Michael Siemer, and Adrien Verdelhan. "International Risk Cycles." *Journal of International Economics* 89, no. 2 (2013): 471-84. doi:10.1016/j.jinteco.2011.10.001.

Horvath, Jaroslav. "Macroeconomic Disasters and the Equity Premium Puzzle: Are Emerging Countries Riskier?" *Journal of Economic Dynamics and Control* 112 (2020): 103852. doi:10.1016/j.jedc.2020.103852.

Kaminsky, Graciela. "Is There a Peso Problem? Evidence from the Dollar/Pound Exchange Rate, 1976-1987." The American Economic Review 83, no. 3 (1993): 450-72.

Lothian, James R., and Liuren Wu. "Uncovered Interest-rate Parity over the past Two Centuries." *Journal of International Money and Finance* 30, no. 3 (2011): 448-73. doi:10.1016/j.jimonfin.2011.01.005.

Lustig, Hanno, Nikolai Roussanov, and Adrien Verdelhan. "Common Risk Factors in Currency Markets." *Review of Financial Studies* 24, no. 11 (2011): 3731-777. doi:10.1093/rfs/hhr068.

Meese, Richard A., and Kenneth Rogoff. "Empirical Exchange Rate Models of the Seventies." *Journal of International Economics* 14, no. 1-2 (1983): 3-24. doi:10.1016/0022-1996(83)90017-x.

Nakamura, Emi, Jón Steinsson, Robert Barro, and José Ursúa. "Crises and Recoveries in an Empirical Model of Consumption Disasters." *American Economic Journal: Macroeconomics* 5, no. 3 (2013): 35-74. doi:10.1257/mac.5.3.35.

Rietz, Thomas A. "The Equity Risk Premium a Solution." *Journal of Monetary Economics* 22, no. 1 (1988): 117-31. doi:10.1016/0304-3932(88)90172-9.

Wachter, Jessica A. "Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?" *The Journal of Finance* 68, no. 3 (2013): 987-1035. doi:10.1111/jofi.12018.

Appendix A: Model Moment Derivations

State Variable Moments:

Moments of z_{2t}.

$$\boldsymbol{E[z_{2t}] = \delta_2}$$

$$\boldsymbol{Var[z_{2t}] = \frac{{\sigma_2}^2}{1-(\varphi_2)^2}}$$

Moments of z_{1t}.

$$E[z_{1t}] = (1-\varphi_1)\delta_1 + \varphi_1 E[z_{1t}] - (E[d_{t+1}])$$

$$= (1-\varphi_1)\delta_1 + \varphi_1 E[z_{1t}] - \theta\omega(E[z_{2t}])$$

$$\boldsymbol{E[z_{1t}] = \delta_1 - \frac{\theta\omega\delta_2}{1-\varphi_1}}$$

$$Var[z_{1t}] = (\varphi_1)^2 Var[z_{1t}] + (\sigma_1)^2 + Var[d_{t+1}] - 2\varphi_1 cov[z_{1t}, d_{t+1}]$$

$$Var[z_{1t}] = \frac{(\sigma_1)^2 + Var[d_{t+1}] - 2\varphi_1 cov[z_{1t}, d_{t+1}]}{1-(\varphi_1)^2}$$

$$Var[d_{t+1}] = Var\left[E_{j_{t+1}}[d_{t+1}]\right] + E\left[Var_{j_{t+1}}[d_{t+1}]\right]$$

$$= Var\left[E_{j_{t+1}}[d_{t+1}]\right] + E\left[Var_{j_{t+1}}[d_{t+1}]\right]$$

$$= Var[\theta j_{t+1}] + E[\rho^2 j_{t+1}]$$

$$= Var\left[E_{z_{2t}}[\theta j_{t+1}]\right] + E\left[Var_{z_{2t}}[\theta j_{t+1}]\right] + E\left[E_{z_{2t}}[\rho^2 j_{t+1}]\right]$$

$$Var[d_{t+1}] = \theta^2\omega^2\left(\frac{{\sigma_2}^2}{1-(\varphi_2)^2}\right) + \theta^2\omega\delta_2 + \rho^2\omega\delta_2$$

$$cov[z_{1t}, d_{t+1}] = cov\left[E_t[z_{1t}], E_t[d_{t+1}]\right] + E\left[cov_t[z_{1t}, d_{t+1}]\right]$$

$$cov[z_{1t}, d_{t+1}] = cov\left[z_{1t}, E_t\left[E_{t,j_{t+1}}[d_{t+1}]\right]\right] + 0$$

$$= cov\left[z_{1t}, E_t[\theta j_{t+1}]\right]$$

$$= cov[z_{1t}, \theta\omega z_{2t}]$$

$$cov[z_{1t}, d_{t+1}] = \theta\omega cov[z_{1t}, z_{2t}]$$

$$\boldsymbol{Var[z_{1t}] = \frac{1}{1-(\varphi_1)^2}\left[(\sigma_1)^2 + \theta^2\omega^2\left(\frac{{\sigma_2}^2}{1-(\varphi_2)^2}\right) + \theta^2\omega\delta_2 + \rho^2\omega\delta_2 - 2\varphi_1\theta\omega cov[z_{1t}, z_{2t}]\right]}$$

Covariance of z_{1t} and z_{2t}.

$$cov[z_{1t+1}, z_{2t+1}] = cov\left[E_{t,d_{t+1}}[z_{1t+1}], E_{t,d_{t+1}}[z_{2t+1}]\right] + E\left[cov_{t,d_{t+1}}[z_{1t+1}, z_{2t+1}]\right]$$

$$E\left[cov_{t,d_{t+1}}[z_{1t+1}, z_{2t+1}]\right] = E\left[cov[\sigma_1\varepsilon_{1t+1}, \sigma_2\varepsilon_{2t+1}]\right] = 0$$

$$cov\left[E_{t,d_{t+1}}[z_{1t+1}], E_{t,d_{t+1}}[z_{2t+1}]\right] = cov[\varphi_1 z_{1t} + \sigma_1\varepsilon_{1t+1} - d_{t+1}\ ,\ \varphi_2 z_{2t} + \sigma_2\varepsilon_{2t+1}]$$

$$= \varphi_1\varphi_2 cov[z_{1t}, z_{2t}] + \varphi_1\sigma_2 cov[z_{1t}, \varepsilon_{2t+1}] + \sigma_1\varphi_2 cov[\varepsilon_{1t+1}, z_{2t}] + \sigma_1\sigma_2 cov[\varepsilon_{1t+1}, \varepsilon_{2t+1}] - \varphi_2 cov[d_{t+1}, z_{2t}] - \sigma_2 cov[d_{t+1}, \varepsilon_{2t+1}]$$

$$cov\left[E_{t,d_{t+1}}[z_{1t+1}], E_{t,d_{t+1}}[z_{2t+1}]\right] = \varphi_1\varphi_2 cov[z_{1t}, z_{2t}] - \varphi_2 cov[d_{t+1}, z_{2t}]$$

$$cov[d_{t+1}, z_{2t}] = E[d_{t+1}z_{2t}] - E[d_{t+1}]E[z_{2t}]$$

$$E[d_{t+1}z_{2t}] = E\big[E_t[d_{t+1}z_{2t}]\big]$$

$$= E[E_t[d_{t+1}]z_{2t}]$$

$$= E\left[E_t\left[E_{t,j_{t+1}}[d_{t+1}]\right]z_{2t}\right]$$

$$= E[E_t[\theta\omega z_{2t}]z_{2t}]$$

$$= \theta\omega E[(z_{2t})^2]$$

$$E[d_{t+1}]E[z_{2t}] = \theta\omega(E[z_{2t}])^2$$

$$cov[d_{t+1}, z_{2t}] = (\theta\omega)Var[z_{2t}]$$

$$cov[z_{1t+1}, z_{2t+1}] = \varphi_1\varphi_2 cov[z_{1t}, z_{2t}] - \varphi_2(\theta\omega)Var[z_{2t}]$$

$$\boldsymbol{cov[z_{1t}, z_{2t}] = -\frac{\varphi_2(\theta\omega)Var[z_{2t}]}{1-\varphi_1\varphi_2}}$$

Interest Rate Differential Moments:

$$r_{1t}^q - r_{1t}^i = f_t - s_t = \left(a^i - a^q\right) + \left(\gamma_1^i - \gamma_1^q\right)z_{1t} + \left(\gamma_2^i - \gamma_2^q\right)z_{2t} + \frac{\left(\lambda_1^i\right)^2 - \left(\lambda_1^q\right)^2 + \left(\lambda_3^i\right)^2 - \left(\lambda_3^q\right)^2}{2} + \left(e^{\left[-\lambda_2^i\theta + \left(\lambda_2^i\right)^2\rho^2/2\right]} - e^{\left[-\lambda_2^q\theta + \left(\lambda_2^q\right)^2\rho^2/2\right]}\right)\omega z_{2t}$$

Expectation of interest rate differential:

$$E\left[r_{1t}^q - r_{1t}^i\right] = \left(a^i - a^q\right) + \left(\gamma_1^i - \gamma_1^q\right)E[z_{1t}] + \left(\gamma_2^i - \gamma_2^q\right)E[z_{2t}] + \frac{\left(\lambda_1^i\right)^2 - \left(\lambda_1^q\right)^2 + \left(\lambda_3^i\right)^2 - \left(\lambda_3^q\right)^2}{2} + \left(e^{\left[-\lambda_2^i\theta + \left(\lambda_2^i\right)^2\rho^2/2\right]} - e^{\left[-\lambda_2^q\theta + \left(\lambda_2^q\right)^2\rho^2/2\right]}\right)\omega E[z_{2t}]$$

Variance of interest rate differential:

$$Var\left[r_{1t}^q - r_{1t}^i\right] = E\left[Var_t\left[r_{1t}^q - r_{1t}^i\right]\right] + Var\left[E_t\left[r_{1t}^q - r_{1t}^i\right]\right]$$

$$= 0 + Var\left[E_t\left[r_{1t}^q - r_{1t}^i\right]\right]$$

$$Var\left[r_{1t}^q - r_{1t}^i\right] = \left(\gamma_1^i - \gamma_1^q\right)^2 Var[z_{1t}] + \left[\left(\gamma_2^i - \gamma_2^q\right) + \left(e^{\left[-\lambda_2^i\theta + \left(\lambda_2^i\right)^2\rho^2/2\right]} - e^{\left[-\lambda_2^q\theta + \left(\lambda_2^q\right)^2\rho^2/2\right]}\right)\omega\right]^2 Var[z_{2t}] + 2\left(\gamma_1^i - \gamma_1^q\right)\left[\left(\gamma_2^i - \gamma_2^q\right) + \left(e^{\left[-\lambda_2^i\theta + \left(\lambda_2^i\right)^2\rho^2/2\right]} - e^{\left[-\lambda_2^q\theta + \left(\lambda_2^q\right)^2\rho^2/2\right]}\right)\omega\right] cov[z_{1t}, z_{2t}]$$

1st order autocovariance of interest rate differential:

$$cov\left[\left(r_{1t}^{q}-r_{1t}^{i}\right),\left(r_{1t-\tau}^{q}-r_{1t-\tau}^{i}\right)\right]=cov\left[E_{t-\tau}\left[r_{1t}^{q}-r_{1t}^{i}\right],E_{t-\tau}\left[r_{1t-\tau}^{q}-r_{1t-\tau}^{i}\right]\right]+$$
$$E\left[cov_{t-\tau}\left[\left(r_{1t}^{q}-r_{1t}^{i}\right),\left(r_{1t-\tau}^{q}-r_{1t-\tau}^{i}\right)\right]\right]$$

$$E\left[cov_{t-\tau}\left[\left(r_{1t}^{q}-r_{1t}^{i}\right),\left(r_{1t-\tau}^{q}-r_{1t-\tau}^{i}\right)\right]\right]=0$$

$$cov\left[E_{t-\tau}\left[r_{1t}^{q}-r_{1t}^{i}\right],E_{t-\tau}\left[r_{1t-\tau}^{q}-r_{1t-\tau}^{i}\right]\right]$$
$$=\ cov\left[E_{t-\tau}[az_{1t}+bz_{2t}],E_{t-\tau}[az_{1t-\tau}+bz_{2t-\tau}]\right]$$

$$E_{t-\tau}[az_{1t}+bz_{2t}]=aE_{t-\tau}[z_{1t}]+bE_{t-\tau}[bz_{2t}]$$

$$E_{t-\tau}[z_{1t}]=E_{t-\tau}[(1-\varphi_1)\delta_1+\varphi_1 z_{1t-1}+\sigma_1\varepsilon_{1t}-d_t]$$

$$=E_{t-\tau}\left[(1-\varphi_1)\delta_1+\varphi_1\left((1-\varphi_1)\delta_1+\varphi_1 z_{1t-2}+\sigma_1\varepsilon_{1t-1}-d_{t-1}\right)+\sigma_1\varepsilon_{1t}-d_t\right]$$

$$=E_{t-\tau}[(1+\varphi_1)(1-\varphi_1)\delta_1+(\varphi_1)^2 z_{1t-2}+\varphi_1\sigma_1\varepsilon_{1t-1}-\varphi_1 d_{t-1}+\sigma_1\varepsilon_{1t}-d_t]$$

...

$$E_{t-\tau}[z_{1t}]=E_{t-\tau}\left[(1-\varphi_1)\delta_1\sum_{n=0}^{\tau-1}(\varphi_1)^n+(\varphi_1)^\tau z_{1t-\tau}-\sum_{n=0}^{\tau-1}(\varphi_1)^n(d_{t-n})\right]$$
$$=(1-\varphi_1)\delta_1\sum_{n=0}^{\tau-1}(\varphi_1)^n+(\varphi_1)^\tau z_{1t-\tau}-E_{t-\tau}\left[\sum_{n=0}^{\tau-1}(\varphi_1)^n(d_{t-n})\right]$$

$$E_{t-\tau}[z_{2t}]=E_{t-\tau}\left[(1-\varphi_2)\delta_1\sum_{n=0}^{\tau}(\varphi_2)^n+(\varphi_2)^\tau z_{2t-\tau}\right]$$

Dropping non-variant terms:

$$E_{t-\tau}\left[\sum_{n=0}^{\tau-1}(\varphi_1)^n(d_{t-n})\right]=\sum_{n=0}^{\tau-1}(\varphi_1)^n E_{t-\tau}[d_{t-n}]=\omega\theta\sum_{n=0}^{\tau-1}(\varphi_1)^n(\varphi_2)^{\tau-1-n}z_{2t-\tau}$$

$$\gg E_{t-\tau}[z_{1t}]=(\varphi_1)^\tau z_{1t-\tau}+\omega\theta(z_{2t-\tau})\sum_{n=0}^{\tau-1}(\varphi_1)^n(\varphi_2)^{\tau-1-n}$$

$$\gg E_{t-\tau}[z_{2t}]=(\varphi_2)^\tau z_{2t-\tau}$$

$$cov\left[E_{t-\tau}[az_{1t}+bz_{2t}],E_{t-\tau}[az_{1t-\tau}+bz_{2t-\tau}]\right]=$$

$$cov\left[a\left((\varphi_1)^\tau z_{1t-\tau}+\omega\theta(z_{2t-\tau})\sum_{n=0}^{\tau-1}(\varphi_1)^n(\varphi_2)^{\tau-1-n}\right)+b((\varphi_2)^\tau z_{2t-\tau})\ ,az_{1t-\tau}+bz_{2t-\tau}\right]$$

$$= (\varphi_1)^\tau a^2 Var[z_{1t}] + b\left[b(\varphi_2)^\tau + \left(a\omega\theta\sum_{n=0}^{\tau-1}(\varphi_1)^n(\varphi_2)^{\tau-1-n}\right)\right]Var[z_{2t}]$$

$$+ a\left[b(\varphi_1)^\tau + b(\varphi_2)^\tau + \left(a\omega\theta\sum_{n=0}^{\tau-1}(\varphi_1)^n(\varphi_2)^{\tau-1-n}\right)\right]cov[z_{1t}, z_{2t}]$$

$\gg cov\left[\left(r_{1t}^q - r_{1t}^i\right), \left(r_{1t-\tau}^q - r_{1t-\tau}^i\right)\right] =$

$(\varphi_1)^\tau a^2 Var[z_{1t}] + b[b(\varphi_2)^\tau + (a\omega\theta\sum_{n=0}^{\tau-1}(\varphi_1)^n(\varphi_2)^{\tau-1-n})]Var[z_{2t}] + a[b(\varphi_1)^\tau + b(\varphi_2)^\tau + (a\omega\theta\sum_{n=0}^{\tau-1}(\varphi_1)^n(\varphi_2)^{\tau-1-n})]cov[z_{1t}, z_{2t}]$

$Where\ a = (\gamma_1^i - \gamma_1^q)\ and\ b = \left[(\gamma_2^i - \gamma_2^q) + (e^{\left[-\lambda_2^i\theta + \left(\lambda_2^i\right)^2\rho^2/2\right]} - e^{\left[-\lambda_2^q\theta + \left(\lambda_2^q\right)^2\rho^2/2\right]})\omega\right]$

Exchange Rate Moments:

$s_{t+1} - s_t = (a^i - a^q) + (\gamma_1^i - \gamma_1^q)z_{1t} + (\gamma_2^i - \gamma_2^q)z_{2t} + (\lambda_1^i - \lambda_1^q)\varepsilon_{1t+1} - (\lambda_2^i - \lambda_2^q)d_{t+1} + (\lambda_3^i - \lambda_3^q)\varepsilon_{2t+1}$

Expectation of log exchange rate:

$E[s_{t+1} - s_t] = E\left[E_t[s_{t+1} - s_t]\right]$

$E[s_{t+1} - s_t] = (a^i - a^q) + (\gamma_1^i - \gamma_1^q)E[z_{1t}] + (\gamma_2^i - \gamma_2^q)E[z_{2t}] - (\lambda_2^i - \lambda_2^q)\theta\omega E[z_{2t}]$

Variance of log exchange rate:

$Var[s_{t+1} - s_t] = E\left[Var_t[s_{t+1} - s_t]\right] + Var\left[E_t[s_{t+1} - s_t]\right]$

$$E\left[Var_t[s_{t+1} - s_t]\right] = E\left[\left(\lambda_1^i - \lambda_1^q\right)^2 + \left(\lambda_3^i - \lambda_3^q\right)^2 + \left(\lambda_2^i - \lambda_2^q\right)^2 Var_t[d_{t+1}]\right]$$

$$Var_t[d_{t+1}] = E_t\left[Var_{t,j_{t+1}}[d_{t+1}]\right] + Var_t\left[E_{t,j_{t+1}}[d_{t+1}]\right]$$

$$Var_t[d_{t+1}] = E_t[\rho^2 j_{t+1}] + Var_t[\theta j_{t+1}]$$

$$Var_t[d_{t+1}] = \rho^2\omega z_{2t} + \theta^2\omega z_{2t}$$

$$E\left[Var_t[s_{t+1} - s_t]\right] = \left(\lambda_1^i - \lambda_1^q\right)^2 + \left(\lambda_3^i - \lambda_3^q\right)^2 + \left(\lambda_2^i - \lambda_2^q\right)^2(\rho^2 + \theta^2)\omega E[z_{2t}]$$

$$Var\left[E_t[s_{t+1} - s_t]\right] = Var\left[(a^i - a^q) + (\gamma_1^i - \gamma_1^q)z_{1t} + (\gamma_2^i - \gamma_2^q)z_{2t} - (\lambda_2^i - \lambda_2^q)\theta\omega z_{2t}\right]$$

$$Var\left[E_t[s_{t+1} - s_t]\right] = \left(\gamma_1^i - \gamma_1^q\right)^2 Var[z_{1t}] + \left((\gamma_2^i - \gamma_2^q) - (\lambda_2^i - \lambda_2^q)\theta\omega\right)^2 Var[z_{2t}] + 2\left((\gamma_2^i - \gamma_2^q) - (\lambda_2^i - \lambda_2^q)\theta\omega\right)(\gamma_1^i - \gamma_1^q)cov[z_{1t}, z_{2t}]$$

$Var[s_{t+1} - s_t]$

$$= \left(\lambda_1^i - \lambda_1^q\right)^2 + \left(\lambda_3^i - \lambda_3^q\right)^2 + \left(\lambda_2^i - \lambda_2^q\right)^2 (\rho^2 + \theta^2)\omega E[z_{2t}]$$
$$+ \left(\gamma_1^i - \gamma_1^q\right)^2 Var[z_{1t}] + \left(\left(\gamma_2^i - \gamma_2^q\right) - \left(\lambda_2^i - \lambda_2^q\right)\theta\omega\right)^2 Var[z_{2t}]$$
$$+ 2\left(\left(\gamma_2^i - \gamma_2^q\right) - \left(\lambda_2^i - \lambda_2^q\right)\theta\omega\right)\left(\gamma_1^i - \gamma_1^q\right) cov[z_{1t}, z_{2t}]$$

Appendix B: GMM

To simplify notation, stack the variables and the unknown parameters into vectors, $z_t = [x_t\ y_t]^{\mathrm{T}}\ and\ \Psi_0 = \left[\varphi_2\ \sigma_2\ \lambda_2^i\theta\ \lambda_2^q\theta\ \left(\gamma_2^i - \gamma_2^q\right)\right]^{\mathrm{T}}$. Other feasible values of parameters will be denoted by Ψ. We can then translate the moment restrictions so that they are linear in expectation and stack them in the following vector valued function:

$$f(z_t, \Psi_0) = \begin{bmatrix} f(z_t, \Psi_0)_1 \\ f(z_t, \Psi_0)_2 \\ f(z_t, \Psi_0)_3 \\ f(z_t, \Psi_0)_4 \\ f(z_t, \Psi_0)_5 \end{bmatrix}$$

Where:

$$f(z_t, \Psi_0)_1 =$$

$$x_t - \left(a^i - a^q\right) - \left(\gamma_2^i - \gamma_2^q + e^{\left[-\lambda_2^i\theta\right]} - e^{\left[-\lambda_2^q\theta\right]}\right)\delta_2$$

$$f(z_t, \Psi_0)_2 =$$

$$x_t{}^2 - \left[\left(a^i - a^q\right) + \left(\gamma_2^i - \gamma_2^q + e^{\left[-\lambda_2^i\theta\right]} - e^{\left[-\lambda_2^q\theta\right]}\right)\delta_2\right]^2$$
$$- \left[\left(\gamma_2^i - \gamma_2^q\right) + \left(e^{-\lambda_2^i\theta} - e^{-\lambda_2^q\theta}\right)\right]^2 \frac{\sigma_2{}^2}{1 - (\varphi_2)^2}$$

$$f(z_t, \Psi_0)_3 =$$

$$x_t x_{t-1} - \left[\left(a^i - a^q\right) + \left(\gamma_2^i - \gamma_2^q + e^{\left[-\lambda_2^i\theta\right]} - e^{\left[-\lambda_2^q\theta\right]}\right)\delta_2\right]^2$$
$$- \varphi_2\left[\left(\gamma_2^i - \gamma_2^q\right) + \left(e^{-\lambda_2^i\theta} - e^{-\lambda_2^q\theta}\right)\right]^2 \frac{\sigma_2{}^2}{1 - (\varphi_2)^2}$$

$$f(z_t, \Psi_0)_4 =$$

$$y_t - \left(a^i - a^q\right) - \left(\gamma_2^i - \gamma_2^q - \lambda_2^i\theta + \lambda_2^q\theta\right)\delta_2$$

$$f(z_t, \Psi_0)_5 =$$

$$y_t{}^2 - \left[\left(a^i - a^q\right) + \left(\gamma_2^i - \gamma_2^q - \lambda_2^i\theta + \lambda_2^q\theta\right)\delta_2\right]^2 - \left(\lambda_2^i\theta - \lambda_2^q\theta\right)^2\delta_2 + \left(\gamma_2^i - \gamma_2^q - \left(\lambda_2^i\theta - \lambda_2^q\theta\right)\right)^2 \frac{\sigma_2{}^2}{1-(\varphi_2)^2}$$

Therefore, the model implies that $E[f(z_t, \Psi_0)] = 0$. Given a sample $\{\boldsymbol{z_0}, \boldsymbol{z_1}, \boldsymbol{z_2}, \dots, \boldsymbol{z_T}\}$, the just-identified GMM estimator of Ψ_0 is denoted as Ψ_T and solves the system of nonlinear equations:

$$\frac{1}{T}\sum_{t=1}^{T} f(z_t, \Psi_T) = 0$$

We now turn to estimating the sampling distribution of Ψ_T. Consider the (5×5) matrix of partial derivatives:

$$\frac{\partial f(z_t,\Psi)}{\partial \Psi} = \begin{bmatrix} \frac{\partial f(z_t,\Psi)_1}{\partial \varphi_2} & \frac{\partial f(z_t,\Psi)_2}{\partial \varphi_2} & \cdots & \frac{\partial f(z_t,\Psi)_5}{\partial \varphi_2} \\ \frac{\partial f(z_t,\Psi)_1}{\partial \sigma_2} & \frac{\partial f(z_t,\Psi)_2}{\partial \sigma_2} & \cdots & \frac{\partial f(z_t,\Psi)_5}{\partial \sigma_2} \\ \frac{\partial f(z_t,\Psi)_1}{\partial \left(\lambda_2^i\theta\right)} & \frac{\partial f(z_t,\Psi)_2}{\partial \left(\lambda_2^i\theta\right)} & \cdots & \frac{\partial f(z_t,\Psi)_5}{\partial \left(\lambda_2^i\theta\right)} \\ \frac{\partial f(z_t,\Psi)_1}{\partial \left(\lambda_2^q\theta\right)} & \frac{\partial f(z_t,\Psi)_2}{\partial \left(\lambda_2^q\theta\right)} & \cdots & \frac{\partial f(z_t,\Psi)_5}{\partial \left(\lambda_2^q\theta\right)} \\ \frac{\partial f(z_t,\Psi)_1}{\partial \left(\gamma_2^i - \gamma_2^q\right)} & \frac{\partial f(z_t,\Psi)_2}{\partial \left(\gamma_2^i - \gamma_2^q\right)} & \cdots & \frac{\partial f(z_t,\Psi)_5}{\partial \left(\gamma_2^i - \gamma_2^q\right)} \end{bmatrix}$$

Given the separable structure of moment restrictions, this partial derivative matrix will be the same for each t:

$$\frac{1}{T}\sum_{t=1}^{T} \left.\frac{\partial f(z_t,\Psi)}{\partial \Psi}\right|_{\Psi=\Psi_T} = \left.\frac{\partial f(z_t,\Psi)}{\partial \Psi}\right|_{\Psi=\Psi_T}$$

Denote the evaluation of this partial derivative at π_T as G_T:

$$G_T = \left.\frac{\partial f(z_t,\Psi)}{\partial \Psi}\right|_{\Psi=\Psi_T}$$

The other matrix we need is:

$$V_T = Var\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} f(z_t, \Psi_0)\right]$$

The Newey-West estimator is:

$$\hat{V}_T = \Omega_{T0} + \sum_{j=1}^{12} w(j,n)\left[\Omega_{Tj} + \Omega_{Tj}^{\mathrm{T}}\right]$$

Where Ω_{Tj} is defined as:

$$\Omega_{Tj} = \frac{1}{\sqrt{T}} \sum_{t=j+1}^{T} [f(z_t, \Psi_T) f(z_t, \Psi_T)^T]$$

And the weighting function as follows:

$$w(j,n) = 1 - \frac{j}{n+1}$$

Given $\hat{V}_T$ and G_T, the asymptotic covariance matrix for π_T can be calculated as follows:

$$H_T = \frac{1}{T}\left(G_T^{-1}\hat{V}_T[G_T^{\mathrm{T}}]^{-1}\right)$$

Denoting h_{iT} as the ith diagonal element of H_T and Ψ_{iT} as the ith element of π_T. Then, in large samples, $\sqrt{h_{iT}}$ is the standard error of Ψ_{iT}.

To calculate the standard error of $\frac{\lambda_2^i}{\lambda_2^q} - 1$, we consider the function:

$$g(x,y) = \frac{x}{y} - 1$$

Denote the two vectors composed of the 3rd and the 4th elements of Ψ_T and Ψ_0 respectively, as $\boldsymbol{\lambda_T}$ and $\boldsymbol{\lambda_0}$**.** We can take the mean-value expansion of g around $\boldsymbol{\lambda_0}$,

$$g(\boldsymbol{\lambda_T}) = g(\boldsymbol{\lambda_0}) + \nabla g(\boldsymbol{c})(\boldsymbol{\lambda_T} - \boldsymbol{\lambda_0}),$$

where $\boldsymbol{c}$ lies between $\boldsymbol{\lambda_T}$ and $\boldsymbol{\lambda_0}$ and $\nabla g(\boldsymbol{c})$ is the (1×2) gradient of g evaluated at $\boldsymbol{c}$. By the law of large numbers, $\boldsymbol{\lambda_T} \to \boldsymbol{\lambda_0}$ and $\boldsymbol{c} = \boldsymbol{\lambda_T}$. Therefore, the standard error of $g(\boldsymbol{\lambda_T})$ is:

$$\left[\left(\nabla g(\boldsymbol{\lambda_T})\right) H_T^{\lambda} \left(\nabla g(\boldsymbol{\lambda_T})\right)^{\mathrm{T}}\right]^{1/2},$$

where H_T^{λ} is the (2×2) sub-matrix of H_T:

$$H_T^{\lambda} = \begin{bmatrix} H_{T(3,3)} & H_{T(3,4)} \\ H_{T(4,3)} & H_{T(4,4)} \end{bmatrix}$$