Online Appendix for "Bond Convenience Yields and Exchange Rate Dynamics

Rosen Valchev (Boston College)

A Data Description

A.1 Exchange Rates

The data set consists of forward and spot exchange rates from Reuters/WMR and Barclays, and is available on Datastream. It includes the Euro and the currencies of the following 18 advanced OECD countries: Australia, Austria, Belgium, Canada, Denmark, France, Germany, Ireland, Italy, Japan, the Netherlands, Norway, New Zealand, Portugal, Spain, Sweden, Switzerland and the UK.

The data spans the time period 1976:M1-2013:M6 and is at a daily frequency. The data on the Euro-legacy currencies (e.g. France, Austria, etc.), except for the German Deutsch Mark (DEM), ends in December 1998. As is common in the literature, instead of including separate DEM and EUR series, I combine the two by appending the Euro to the end of the DEM series. This creates a single long series that spans the whole time frame.

The data consists of forward and spot exchange rates, and I construct interest rate differentials from the Covered Interest Parity (CIP):

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

This is the standard practice in the literature because the data on forward contracts is better than data on short-term interest rates, since the forward market is deep and liquid.

A.2 Data for Section 6

The data on government debt and GDP is from the OECD database. The data on Commercial Paper is from the Federal Reserve, Board of Governors. The NFA series is constructed from data on foreign assets and liabilities from the IMF's International Financial Statistics Database. Data on interest rates – AAA, BAA, Treasury Notes and Bills – is from Datastream. The equity volatility measure is built from MSCI Stock Price indices, also obtained from Datastream.

B The UIP Condition

I define S_t to be the exchange rate, in terms of home currency per one unit of foreign currency (e.g. 1.25 USD per EUR), and i_t and i_t^* as the nominal interest rates on default-free bonds at home and abroad. For ease of exposition, I will refer to the US dollar as the "home" currency and the Euro as the "foreign" currency. A \$1 investment in US bonds at time t offers a return of $1 + i_t$ dollars next period. The same \$1 invested in Euro denominated bonds would earn $\frac{S_{t+1}}{S_t}(1 + i_t^*)$ dollars next period. First, we need to exchange this one dollar for Euros and obtain $\frac{1}{S_t}$ EUR in return. Investing this amount of Euros earns a gross interest rate of $1 + i_t^*$ that next period can be exchanged back into dollars at the rate S_{t+1}, for a total return of $\frac{S_{t+1}}{S_t}(1 + i_t^*)$ dollars.

Assuming that the law of one price holds, there exists a stochastic discount factor M_{t+1}, such that

$$E_t(M_{t+1}(1 + i_t)) = 1 \tag{1}$$

$$E_t(M_{t+1}\frac{S_{t+1}}{S_t}(1 + i_t^*)) = 1. \tag{2}$$

A straightforward way to obtain the Uncovered Interest Parity condition is to log-linearize the two equations, subtract them from one another and re-arrange to arrive at

$$E_t(s_{t+1} - s_t + i_t^* - i_t) = 0$$

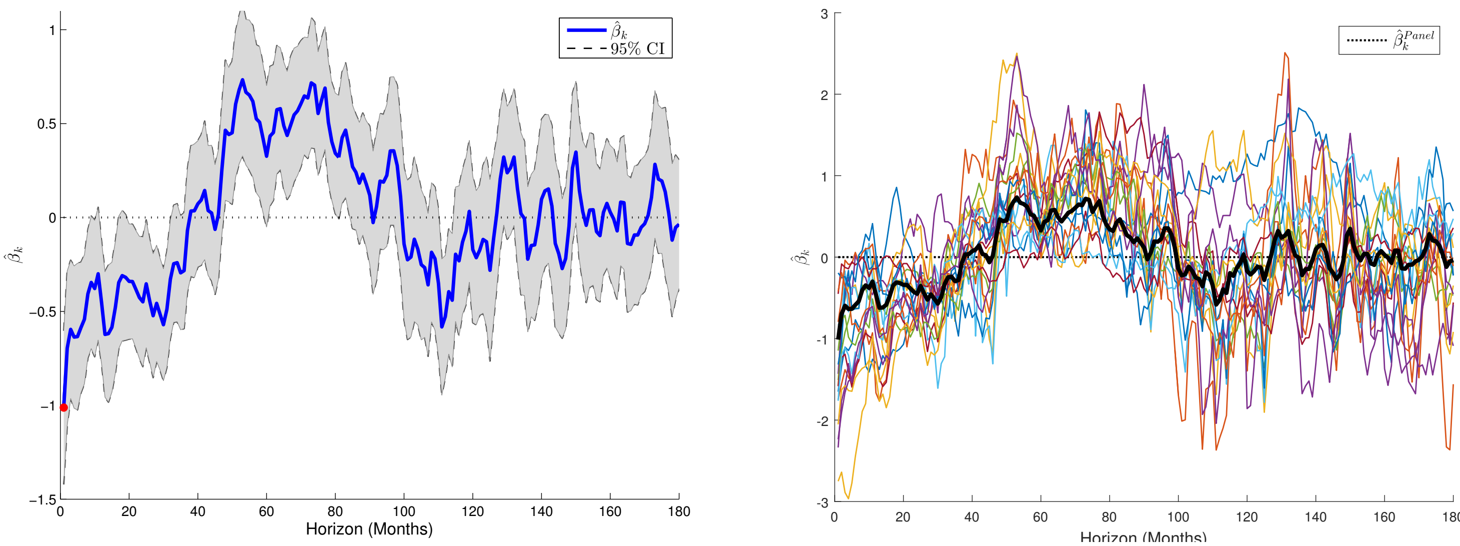
where lower case letters represent variables in logs and I have used the approximation $i_t \approx \ln(1 + i_t)$.[1] Thus, up to a first-order approximation, the expected return on foreign bonds, $E_t(s_{t+1} - s_t + i_t^*)$, equals the expected return on the home bond, i_t. This restricts the joint dynamics of exchange rates and interest rates, and delivers strong implications for exchange rate behavior. The condition obtains in a large class of standard open economy models.

B.1 Including post-crisis data

In this section I re-estimate the main empirical specification, the UIP regressions

$$\lambda_{j,t+k} = \alpha_{j,k} + \beta_k(i_t - i_{j,t}^*) + \varepsilon_{j,t+k}, \tag{3}$$

[1] The log-linearization is typically done around the symmetric steady state where $S_{t+1} = S_t = 1$ and $i_t = i_t^*$, because this allows us to express the condition in terms of the log-variables themselves. But the log-linearized condition holds for any arbitrary point of approximation.



(a) Panel estimates

(b) Currency-by-currency estimates

Figure 1: UIP Regression at horizons from 1 to 180 months

and the exchange rate impulse response

$$s_{j,t+k} - s_{jt} = \alpha_{j,k} + \gamma_k(i_t - i^*_{j,t}) + \varepsilon_{j,t+k}, \tag{4}$$

on the full sample 1976-2013, which includes the financial crisis and the subsequent period.

The resulting estimates are plotted in Figures 1 and 2 and show that there is no ostensible difference from the estimates on the truncated sample that excludes the financial crisis. Thus, the cyclicality of UIP violations is not something that is confined to either sub-sample and suggests that potential structural breaks are not affecting the results. The impulse response of the exchange rate is also virtually identical to the benchmark results, and displays clear non-monotonic dynamics.

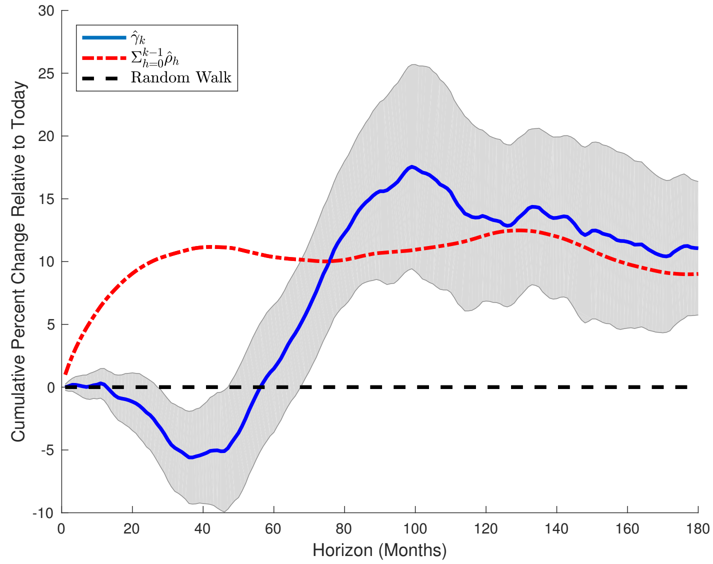
B.2 Exchange Rate Changes Predictability

To complement the discussion in Section 1, here I show the predictability pattern of exchange rate changes, Δs_{t+k+1}, at different horizons. To do so, I estimate the regression

$$\Delta s_{j,t+k} = \alpha_{jk} + \tilde{\gamma}_k(i_t - i^*_{jt}) + \varepsilon_{j,t+k}$$

and plot the coefficients $\tilde{\gamma}_k$. Those coefficients summarize the predictability in the one month exchange rate change at different horizons. For example, $\tilde{\gamma}_1$ captures the predictability of the change between t and $t+1$, and $\tilde{\gamma}_{k+1}$ more generally captures the predictability of the

Figure 2: Exchange Rate IRF



change between time periods $t + k$ and $t + k + 1$.

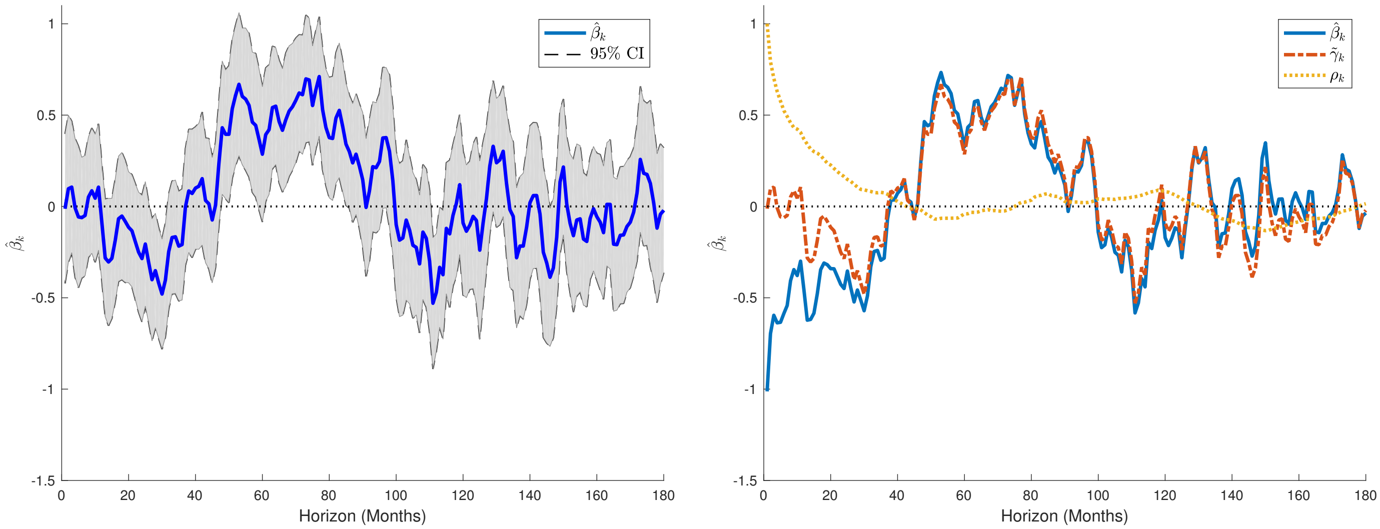
The results are plotted in the left panel of Figure 3. As we would anticipate from the results plotted in Figure 2, we see that there is no exchange rate predictability at short horizons of up to one to one and a half years. Then, at horizons between 18 to 36 months higher current interest rate depreciation forecasts an exchange rate appreciation, and lastly, at horizons between roughly 4 to 7 years, higher interest rate differentials today forecast exchange rate depreciation. Note that the IRF of the level of the exchange rate, γ_k is simply equal to the sum of the coefficients $\tilde{\gamma}_k$ plotted here:

$$\gamma_k = \sum_{i=1}^{k} \tilde{\gamma}_k$$

Moreover, panel (b) on the right plots all three coefficients, the predictability in excess returns (β_k), predictability in exchange rate changes ($\tilde{\gamma}_k$), the impulse response of the interest rate differential (ρ_k) together. Note that the regression coefficient on the currency excess returns is simply the difference of the other two:

$$\beta_k = \tilde{\gamma}_k - \rho_k.$$

So as we can see, the predictability in the excess currency returns at horizons of over 36 months is almost exclusively due to predictability in exchange rate changes. In particular, at these longer horizons the exchange rate is expected to sustain a significant depreciation (positive $\tilde{\gamma}_k$), which results in negative expected excess currency returns at those horizons.



(a) $\tilde{\gamma}_k$ (b) Decompostion

Figure 3: Exchange Rate Changes Predictability

In conclusion, the results of this section confirm that the change in the sign of the excess return predictability (the sign on the β_k coefficients) is driven by a change in the sign of the predictability in high frequency exchange rate movements at longer horizons. This complements the discussion in Section **??** which argues that it is the changing nature of exchange rate predictability that underlies the estimated cyclicality of the currency excess returns.

C Proofs

C.1 LEMMA 1:

LEMMA 1. *The equilibrium is determinate if and only if we have one of two policy regimes:*

(i) Active Monetary, Passive Fiscal policy: $\phi_\pi > 1$, $\kappa_b \in (\theta - \theta_2, \frac{1+\rho_\tau}{1-\rho_\tau}(\theta + \theta_2))$, $\rho_\tau \in [0, \frac{\theta_2}{\theta})$.

(ii) Passive Monetary, Active Fiscal policy: $\phi_\pi < 1$, $\kappa_b \notin (\theta - \theta_2, \frac{1+\rho_\tau}{1-\rho_\tau}(\theta + \theta_2))$, $\rho_\tau \in [0, 1)$.

where $\theta > \theta_2 \geq 1$, *with* $\theta = (1+i)(1+\gamma_\Psi + \gamma_M)$, $\theta_2 = 1 + \gamma_M(1+i)$, *and* $\gamma_\Psi > 0$, *and* $\gamma_M \geq 0$ *are log-linearization constants defined in the Appendix.*

Proof. I will first show the if direction. The equilibrium of the model is described by four (log-linearized) equations: Euler equation for home bonds, government budget, the Taylor rule and the tax rule. These equations determine the dynamics of the four domestic equilibrium variables – inflation, interest rates, government debt and taxes – and represent a closed system that can be solved independent of foreign variables considerations.

Using the fact that consumption and foreign bonds holdings are constant, the the log-linearized MRS becomes,

$$\hat{M}_{t+1} = \gamma_M(\hat{b}_{h,t+1} - \hat{b}_{h,t})$$

where $\gamma_M = \frac{u_{cb_h}(c,b_h,b_f)}{u_c(c,b_h,b_f)} b_h > 0$, and the log-linearized convenience benefit is:

$$\frac{\Psi^H}{\beta(1+i)}\hat{\Psi}^H = -\gamma_\Psi \hat{b}_{ht}$$

where $\gamma_\Psi = -\frac{b_h}{\beta(1+i)} \frac{1}{u_c(c,b_h,b_f)}(u_{b_h b_h}(c, b_h, b_f) - u_{b_h}(c, b_h, b_f)\frac{u_{cb_h}(c,b_h,b_f)}{u_c(c,b_h,b_f)}) > 0$. I am using the convention that $u_x(.)$ represents the partial derivative in respect to x, and $u_{xx}(.)$ represents the second partial and so on. Variables without time subscripts are steady-state values.

Using these relationships, and the fact that in equilibrium home agent bond holdings equal the supply of home government debt, the system of equilibrium conditions becomes

$$\begin{aligned}
\hat{i}_t &= E_t(\hat{\pi}_{t+1}) + \gamma_\Psi \hat{b}_t - \gamma_M(E_t(\hat{b}_{h,t+1}) - \hat{b}_{h,t}) \\
\hat{b}_{ht} + \frac{\tau}{b_h}\hat{\tau}_t &= (1+i)(\hat{b}_{h,t-1} + \hat{i}_{t-1} - \hat{\pi}_t) \\
\hat{i}_t &= \phi_\pi \hat{\pi}_t + v_t \\
\hat{\tau}_t &= \rho_\tau \hat{\tau}_{t-1} + (1-\rho_\tau)\kappa_b \frac{b_h}{\tau}\hat{b}_{h,t-1}
\end{aligned}$$

First, I show that condition (i), Active monetary/passive fiscal policy mix, ensures that a determinate, stable equilibrium exists. Assume that $\phi_\pi > 1$, $\kappa_b \in (\theta - \theta_2, \frac{1+\rho_\tau}{1-\rho_\tau}(\theta + \theta_2))$, and $\rho_\tau \in [0, \frac{\theta_2}{\theta})$, where $\theta = (1+i)(1+\gamma_\Psi+\gamma_M)$, and $\theta_2 = 1 + (1+i)\gamma_M$. Substituting the Taylor rule into the Euler equation for the home bonds, and solving forward for inflation:

$$\begin{aligned}
\hat{\pi}_t &= \frac{1}{\phi_\pi}\Big(E_t(\hat{\pi}_{t+1}) + (\gamma_\Psi + \gamma_M)\hat{b}_t - \gamma_M E_t(\hat{b}_{h,t+1}) - v_t\Big) \\
&\vdots \\
&= \frac{\gamma_M}{\phi_\pi}\hat{b}_{ht} - \frac{v_t}{\phi_\pi} + \frac{\gamma_\Psi + \gamma_M(1-\phi_\pi)}{\phi_\pi}\sum_{j=0}^{\infty}\frac{1}{\phi_\pi^j}E_t(\hat{b}_{h,t+j})
\end{aligned}$$

Next, date the government budget constraint one period forward, take an expectation conditional on time t information and use the Euler equation and the tax rule to substitute out the interest rate and inflation, and arrive at the following 2 equations:

$$E_t \underbrace{\begin{bmatrix} \hat{b}_{h,t+1} \\ \hat{\tau}_{t+1} \end{bmatrix}}_{x_{t+1}} = \underbrace{\begin{bmatrix} \frac{\theta-(1-\rho_\tau)\kappa_b}{\theta_2} & -\frac{\tau}{b}\frac{\rho_\tau}{\theta_2} \\ (1-\rho_\tau)\kappa_b\frac{b}{\tau} & \rho_\tau \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \hat{b}_{ht} \\ \hat{\tau}_t \end{bmatrix}}_{x_t} \tag{5}$$

I will show that condition (i) ensures that the eigenvalues of the auto-regressive matrix A are inside the unit circle, and hence we can use this system to solve for the infinite sum of expected b_{ht} in the expression for equilibrium inflation. The two eigenvalues of A are

$$\lambda_{1,2} = \frac{\theta - (1-\rho_\tau)\kappa_b + \theta_2\rho_\tau \pm \sqrt{(\theta - (1-\rho_\tau)\kappa_b + \theta_2\rho_\tau)^2 - 4\theta\theta_2\rho_\tau}}{2\theta_2}.$$

The eigenvalues are complex conjugates when $(\theta - (1-\rho_\tau)\kappa_b + \theta_2\rho_\tau)^2 - 4\theta\theta_2\rho_\tau < 0$. The left-hand side of this equation defines a quadratic expression in ρ_τ that is convex and crosses zero at the following two points

$$\underline{\rho}(\kappa_b) = \frac{\kappa_b(\kappa_b - \theta) + (\kappa_b + \theta)\theta_2 - 2\sqrt{\kappa_b\theta\theta_2(\kappa_b - \theta + \theta_2)}}{(\theta_2 + \kappa_b)^2}$$

$$\overline{\rho}(\kappa_b) = \frac{\kappa_b(\kappa_b - \theta) + (\kappa_b + \theta)\theta_2 + 2\sqrt{\kappa_b\theta\theta_2(\kappa_b - \theta + \theta_2)}}{(\theta_2 + \kappa_b)^2}$$

Since $\theta_2 < \theta$ it follows that $\underline{\rho}(\kappa_b) < 1$ and since

$$\kappa_b(\kappa_b - \theta) + (\kappa_b + \theta)\theta_2 = \kappa_b(\kappa_b - \theta + \theta_2) + \theta\theta_2$$

it follows that $\underline{\rho}(\kappa_b) > 0$. Moreover, $\underline{\rho}(\kappa_b) \leq \frac{\theta_2}{\theta} \leq \overline{\rho}_\tau(\kappa_b)$, and hence for $\rho_\tau \in [0, \underline{\rho}(\kappa_b)]$ the eigenvalues are real, and for $\rho \in (\underline{\rho}(\kappa_b), \frac{\theta_2}{\theta})$ they are complex conjugates.

First, I address the case where the eigenvalues are complex. Their magnitude its:

$$\begin{aligned} |\lambda_k| &= \frac{1}{2\theta_2}\left((\theta - (1-\rho_\tau)\kappa_b + \theta_2\rho_\tau)^2 + [4\theta\theta_2\rho_\tau - (\theta - (1-\rho_\tau)\kappa_b + \theta_2\rho_\tau)^2]\right)^{\frac{1}{2}} \\ &= \frac{1}{2\theta_2}\sqrt{4\theta\theta_2\rho_\tau} \\ &= \sqrt{\frac{\theta}{\theta_2}\rho_\tau} \end{aligned}$$

and hence $|\lambda_k| < 1$ if and only if $\rho_\tau < \frac{\theta_2}{\theta}$. This is satisfied by condition (i), and hence when the eigenvalues are complex, they lie inside the unit circle.

Next, I address the situation when the eigenvalues are real, $\rho_\tau < \underline{\rho}_\tau(\kappa_b)$. First, I will show that $\kappa_b = \theta - \theta_2$ is the minimum value for which the eigenvalues are both inside the unit circle. For $\kappa_b = \theta - \theta_2$ we have $\underline{\rho}(\kappa_b) = \overline{\rho}(\kappa_b) = \frac{\theta_2}{\theta}$, and hence the roots are real for all values of ρ_τ under condition (i). Moreover, for that value of κ_b:

$$\begin{aligned}
\lambda_1 &= \frac{1}{2\theta_2}(\theta_2 + \rho_\tau\theta + \sqrt{(\theta_2 + \rho_\tau\theta)^2 - 4\theta\theta_2\rho_\tau}) \\
&= \frac{1}{2\theta_2}(\theta_2 + \rho_\tau\theta + \sqrt{(\theta_2 - \rho_\tau\theta)^2}) \\
&= 1
\end{aligned}$$

while $\lambda_2 = \rho_\tau\frac{\theta}{\theta_2} < 1$. Next, notice that when $\kappa_b < \frac{\theta+\theta_2\rho_\tau}{1-\rho_\tau}$ we have $\theta-(1-\rho_\tau)\kappa_b+\theta_2\rho_\tau > 0$ and thus $\lambda_1 > 0$ whenever it is real. Furthermore,

$$\frac{\partial\lambda_1}{\partial\kappa_b} = -\frac{1-\rho_\tau}{2\theta_2} - \frac{(1-\rho_\tau)(\theta - (1-\rho_\tau)\kappa_b + \theta_2\rho_\tau)}{2\theta_2\sqrt{(\theta - (1-\rho_\tau)\kappa_b + \theta_2\rho_\tau)^2 - 4\theta\theta_2\rho_\tau}} < 0$$

and hence for $\kappa_b \in (\theta - \theta_2, \frac{\theta+\theta_2\rho_\tau}{1-\rho_\tau})$ we have $\lambda_1 \in (0, 1)$. Moreover, for those values of κ_b $\lambda_2 > 0$ as well (when real), and since whenever the eigenvalues are real $(\theta - (1-\rho_\tau)\kappa_b + \theta_2\rho_\tau)^2 - 4\theta\theta_2\rho_\tau \geq 0$ and thus $\lambda_2 < \lambda_1$, it follows that

$$0 < \lambda_2 < \lambda_1 < 1$$

for all $\kappa_b \in (\theta - \theta_2, \frac{\theta+\theta_2\rho_\tau}{1-\rho_\tau})$.

On the other hand, if $\kappa_b = \frac{\theta+\theta_2\rho_\tau}{1-\rho_\tau}$, then the eigenvalues are complex for all $\rho_\tau > 0$, and when $\rho_\tau = 0$, then $\lambda_1 = \lambda_2 = 0$.

Lastly, consider $\kappa_b \in (\frac{\theta+\theta_2\rho_\tau}{1-\rho_\tau}, \frac{(\theta+\theta_2)(1+\rho_\tau)}{1-\rho_\tau})$. In this case, whenever the eigenvalues are real they are negative since

$$\begin{aligned}
\lambda_1 &= \frac{\theta - (1-\rho_\tau)\kappa_b + \theta_2\rho_\tau + \sqrt{(\theta - (1-\rho_\tau)\kappa_b + \theta_2\rho_\tau)^2 - 4\theta\theta_2\rho_\tau}}{2} \\
&\leq \frac{\theta - (1-\rho_\tau)\kappa_b + \theta_2\rho_\tau + |\theta - (1-\rho_\tau)\kappa_b + \theta_2\rho_\tau|}{2} \\
&\leq 0
\end{aligned}$$

and thus $\lambda_2 \leq \lambda_1 \leq 0$. Furthermore,

$$\frac{\partial\lambda_2}{\partial\kappa_b} = -\frac{1-\rho_\tau}{2\theta_2} + \frac{(1-\rho_\tau)(\theta - (1-\rho_\tau)\kappa_b + \theta_2\rho_\tau)}{2\theta_2\sqrt{(\theta - (1-\rho_\tau)\kappa_b + \theta_2\rho_\tau)^2 - 4\theta\theta_2\rho_\tau}} < 0$$

since $\theta - (1-\rho_\tau)\kappa_b + \theta_2\rho_\tau < 0$, and at $\kappa_b = \frac{(\theta+\theta_2)(1+\rho_\tau)}{1-\rho_\tau}$ we have

$$\lambda_2 = -1$$

Therefore, for $\kappa_b \in (\theta - \theta_2, \frac{1+\rho_\tau}{1-\rho_\tau}(\theta+\theta_2))$ and $\rho_\tau < \underline{\rho}_\tau(\kappa_b)$ the eigenvalues are real and less than 1 in absolute value. And as we have already shown, since $\rho_\tau < \frac{\theta_2}{\theta}$, whenever the eigenvalues are complex they are also less than 1 in modulus.

Thus, condition (i) implies that the eigenvalues of A lie inside the unit circle, so then

$$\sum_{j=0}^{\infty} \frac{1}{\phi_\pi^j} E_t(\hat{b}_{h,t+j}) = [1,0] * (I - \frac{1}{\phi_\pi} A)^{-1} \begin{bmatrix} \hat{b}_{ht} \\ \hat{\tau}_t \end{bmatrix}$$

and we can use this expression to solve for equilibrium inflation in terms of debt and taxes at time t. We can then substitute the interest rate and inflation, and arrive at a 2 equation system that determines $\hat{b}_{ht}$ and $\hat{\tau}_t$:

$$\begin{bmatrix} \hat{b}_{h,t+1} \\ \hat{\tau}_{t+1} \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{\theta-(1-\rho_\tau)\kappa_b}{\theta_2} & -\frac{\tau}{b}\frac{\rho_\tau}{\theta_2} \\ (1-\rho_\tau)\kappa_b \frac{b}{\tau} & \rho_\tau \end{bmatrix}}_{=A} \begin{bmatrix} \hat{b}_{ht} \\ \hat{\tau}_t \end{bmatrix} + \underbrace{\begin{bmatrix} \frac{1+i}{\phi_\pi} \\ 0 \end{bmatrix}}_{B} v_t \tag{6}$$

Unsurprisingly, the auto-regressive matrix is the same matrix A we have already analyzed. As a result, we know that when condition (i) holds, its eigenvalues are inside the unit circle and we have a stationary solution for debt and taxes.

Now assume that condition (ii) holds so $\phi_\pi < 1$, $\kappa_b \notin (\theta - \theta_2, (\theta+\theta_2)\frac{1+\rho_\tau}{1-\rho_\tau})$, and $\rho_\tau < 1$. In this case we cannot solve for inflation forward, however, equation (5) still holds and now I will show that $\kappa_b \notin (\theta - \theta_2, (\theta+\theta_2)\frac{1+\rho_\tau}{1-\rho_\tau})$ implies that at least one of the eigenvalues of A is greater than 1 in absolute value.

First, note that for $\kappa_b < \theta - \theta_2$

$$(\theta - (1-\rho_\tau)\kappa_b + \rho_\tau\theta_2)^2 - 4\theta\theta_2\rho_\tau \geq 0$$

and hence the eigenvalues are always real. Moreover, above we showed that when the eigenvalues are real, $\frac{\partial\lambda_1}{\partial\kappa_b} < 0$ and that $\lambda_1 = 1$ when $\kappa_b = \theta - \theta_2$, hence it follows that $\lambda_1 > 1$ for any $\kappa_b < \theta - \theta_2$. Similarly, if $\kappa_b > (\theta+\theta_2)\frac{1+\rho_\tau}{1-\rho_\tau}$, the roots are also always real and as we have shown above at $\kappa_b = (\theta+\theta_2)\frac{1+\rho_\tau}{1-\rho_\tau}$, $\lambda_2 = -1$ and it is decreasing in κ_b. So it follows that for $\kappa_b > (\theta+\theta_2)\frac{1+\rho_\tau}{1-\rho_\tau}$, we have $\lambda_2 < -1$, and in either case we have an eigenvalue greater than one.

If A is diagonalizable, we can express equation (5) as

$$E_t(x_{t+1}) = P\Lambda P^{-1}x_t$$

where $x_t = \begin{bmatrix} \hat{b}_{h,t} \\ \hat{\tau}_t \end{bmatrix}$, and Λ is a diagonal matrix with the eigenvalues of A on the diagonal, and P is the matrix of corresponding eigenvectors. We can then multiply on both sides by P^{-1}, define $\tilde{x}_t = P^{-1}x_t$ and obtain the diagonal system

$$E_t(\tilde{x}_{t+1}) = \Lambda\tilde{x}_t$$

and in particular,

$$E_t(\tilde{x}^{(1)}_{t+1}) = \lambda_1\tilde{x}^{(1)}_t \tag{7}$$

where $\tilde{x}^{(1)}_t$ is the first element of the vector. If A is not diagonalizable, then we can use the Jordan Normal form where P is the matrix of generalized eigenvalues, and Λ is upper triangular, with the repeated eigenvalue on the diagonal, and 1 in the upper right corner. We can then use the second equation of the resulting system to arrive at a univariate equation similar to (7) where the repeated eigenvalue $|\lambda| > 1$ is the coefficient. Everything else then follows in the same manner.

We can then solve (7) forward (since $|\lambda_1| > 0$) and obtain

$$\tilde{x}^{(1)}_{t+1} = \lim_{j\to\infty} \frac{1}{\lambda_1^j} E_t(\tilde{x}^{(1)}_{t+j}) = 0$$

Recall that $\tilde{x}_t = P^{-1}x_t$ and hence a linear combination of $\hat{b}_{ht}$ and $\hat{\tau}_t$ is equal to 0, therefore we can write

$$\hat{\tau}_t = K\hat{b}_t$$

for some constant K. Substituting in the tax rule equation for debt, we obtain

$$\hat{\tau}_t = (\rho_\tau - (1-\rho_\tau)\kappa_b \frac{b_h}{\tau} K)\hat{\tau}_{t-1}$$

which implies that the solution is

$$\hat{\tau}_t = \hat{b}_{ht} = 0$$

Next, we can substitute this result in the government budget and obtain the relationship

$$i_{t-1} = \pi_t$$

Substituting in the Taylor rule we find the solution for inflation:

$$\pi_t = \phi_\pi \pi_{t-1} + v_{t-1}$$

Since $\phi_\pi < 1$, this is stationary and this concludes the forward direction of the proof. We have shown that when either conditions (i) or (ii) are satisfied, there is a determinate stationary equilibrium.

In proving the necessary direction, I start with the case where $\phi_\pi > 1$. This time I will first deal with the conditions on κ_b, and to this end assume that $\kappa_b < \theta - \theta_2$. Above we showed that in this case the roots are always real, and that $\lambda_1\Big|_{\kappa_b=\theta-\theta_2} = 1$, and that $\frac{\partial\lambda_1}{\partial\kappa_b} < 0$ for $\kappa < \frac{\theta+\theta_2}{1-\rho_\tau}$ which holds since $\theta - \theta_2 < \frac{\theta+\theta_2}{1-\rho_\tau}$. Therefore, it is immediate that $\kappa_b < \theta - \theta_2$ leads to a root bigger than one and thus explosive solutions.

On the other hand if $\kappa_b > \frac{(\theta+\theta_2)(1+\rho_\tau)}{1-\rho_\tau}$, then

$$(\theta - (1-\rho_\tau)\kappa_b + \rho_\tau\theta_2)^2 - 4\theta\theta_2\rho_\tau \geq 0$$

so the roots are again always real. Moreover, we have already shown that $\lambda_2\Big|_{\kappa_b=\frac{(\theta-\theta_2)(1+\rho_\tau)}{1-\rho_\tau}} = -1$, and that $\frac{\partial\lambda_2}{\partial\kappa_b} < 0$ for $\kappa_b > \frac{(\theta-\theta_2)(1+\rho_{tau})}{1-\rho_\tau}$, hence it follows that $|\lambda_2| > 1$ for all $\kappa_b > \frac{(\theta-\theta_2)(1+\rho_\tau)}{1-\rho_\tau}$, and thus we again have an explosive root.

Next, turn attention to $\rho_\tau > \frac{\theta_2}{\theta}$ and $\kappa_b \in (\theta - \theta_2, \frac{(\theta-\theta_2)(1+\rho_{tau})}{1-\rho_\tau})$. If $\rho_\tau \in [\frac{\theta_2}{\theta}, \overline{\rho}_\tau(\kappa_b))$ then the resulting complex eigenvalues will be outside of the unit circle and there are no non-explosive solutions for debt and taxes. On the other hand, if $\rho_\tau \geq \overline{\rho}_\tau(\kappa_b)$, then

$$\frac{\partial\lambda_1}{\partial\rho_\tau} = \frac{\kappa_b + \theta_2}{2\theta_2} + \frac{1}{2\theta_2}\frac{(\kappa_b+\theta_2)(\theta - (1-\rho_\tau)\kappa_b + \rho_\tau\theta_2) - 2\theta\theta_2}{\sqrt{(\theta - (1-\rho_\tau)\kappa_b + \rho_\tau\theta_2)^2 - 4\theta\theta_2\rho_\tau}} > 0$$

since $\kappa_b + \theta_2 > \theta > 1$ and $(\theta - (1-\rho_\tau)\kappa_b + \rho_\tau\theta_2) - 2\theta\theta_2\rho_\tau \geq 0$. Moreover,

$$\begin{aligned}
\lambda_1\Big|_{\rho_\tau=\overline{\rho}_\tau(\kappa_b)} &= \frac{\theta + \sqrt{\kappa_b\frac{\theta}{\theta_2}(\kappa_b - (\theta-\theta_2))}}{\kappa_b+\theta_2} \\
&> \frac{\theta + (\kappa_b - (\theta-\theta_2))}{\kappa_b+\theta_2} \\
&= 1
\end{aligned}$$

where the inequality follows from the fact that $\theta > \theta_2$, and hence $\kappa_b\frac{\theta}{\theta_2} > \kappa_b > \kappa_b - (\theta - \theta_2)$.

Thus, we see that $\lambda_1 > 1$ and hence we again have an explosive root.

Next, I treat the case $\phi_\pi < 1$. If $\kappa_b \in [\theta - \theta_2, \frac{(\theta-\theta_2)(1+\rho_\tau)}{1-\rho_\tau}]$, then either the auto-regressive matrix A has a unit root (unstable solutions), or it has both eigenvalues inside the unit circle. When both roots are inside the unit circle, then conditional on a process for equilibrium inflation, we can solve for debt and taxes backwards. However, in this case we do not have a determinate solution for inflation – in fact there could be many inflation processes that would satisfy the government budget constraint and the Euler equations for bonds. To see this, you let ε^π_{t+1} be the expectational error defined as

$$\hat{\pi}_{t+1} = E_t(\hat{\pi}_{t+1}) + \varepsilon^\pi_{t+1}$$

Using this expression we can again reduce to a system of 2 equations that define a first-order difference system for $\hat{b}_{ht}$ and $\hat{\tau}_t$, with A as the auto-regressive matrix. That defines stationary solutions for debt and taxes, conditional on the expectational error ε^π_{t+1}. Then, we can substitute the Taylor rule in the Euler equation and arrive at

$$\pi_{t+1} = \phi_\pi \hat{\pi}_t + v_t - (\gamma_\Psi + \gamma_M)\hat{b}_{ht} + \gamma_M E_t(\hat{b}_{h,t+1}) - \varepsilon^\pi_{t+1}$$

Since $\phi_\pi < 1$ and $\hat{b}_{ht}$ is stationary, this defines a stationary process for equilibrium inflation. However, the expectational error ε^π_{t+1} is undetermined, and as a result many different processes for inflation satisfy the equilibrium conditions. Thus, with $\phi_\pi < 1$ and $\kappa_b \in (\theta - \theta_2, \frac{(\theta-\theta_2)(1+\rho_{tau})}{1-\rho_\tau})$ the equilibrium is indeterminate. □

C.2 LEMMA 2:

LEMMA 2. *If $\phi_\pi > 1$, $\kappa_b \in (\theta - \theta_2, \frac{\theta+(\theta_2-1)\rho_\tau}{1-\rho_\tau})$, there exists a threshold $\underline{\rho} > 0$ such that*

(i) If $\rho_\tau \in [0, \underline{\rho}]$ the equilibrium system has two real, positive eigenvalues, and

$$a_{bk} > 0 \text{ for } k = 0, 1, 2, 3, \ldots$$

(ii) If $\rho_\tau \in (\underline{\rho}, \frac{\theta_2}{\theta})$ the equilibrium system has a pair of complex conjugate eigenvalues thus

$$a_{bk} = \eta_1 \cos(k\zeta + \eta_2), \text{ for } k = 1, 2, 3, \ldots$$

where η_1 and η_2 are constants given in the Appendix, and $a_{bk} > 0$ for $k \in \{0, 1\}$.

If instead $\phi_\pi < 1$, $\kappa_b \in [0, \theta - \theta_2)$, $\rho_\tau \in [0, 1)$. Then, the system has two real, positive eigenvalues for all $\rho_\tau \in [0, 1)$. Moreover, debt is in fact constant:

$$a_{bk} = 0 \textit{ for } k = 0, 1, 2, 3, \ldots$$

Proof. **Part (i):** The first part follows directly from the proof of Lemma 1: $\rho_\tau \leq \underline{\rho}_\tau(\kappa_b)$ ensures that the eigenvalues are real, and $\kappa_b \in (\theta - \theta_2, \frac{\theta + (\theta_2 - 1)\rho_\tau}{1 - \rho_\tau})$ ensures they are both positive.

To characterize the IRF note that the Wold decomposition of x_t is

$$x_t = Bv_t + ABv_{t-1} + A^2 Bv_{t-2} + \ldots$$

and use the fact that

$$B = \begin{bmatrix} \frac{1+i}{\phi_\pi} \\ 0 \end{bmatrix} v_t$$

to obtain

$$\hat{b}_{ht} = \frac{1+i}{\phi_\pi}(v_t + a_{11}^{(1)} v_{t-1} + a_{11}^{(2)} v_{t-2} + a_{11}^{(3)} v_{t-3} + \ldots)$$

$$\hat{\tau}_t = \frac{1+i}{\phi_\pi}(a_{21}^{(1)} v_{t-1} + a_{21}^{(2)} v_{t-2} + a_{21}^{(3)} v_{t-3} + \ldots)$$

where $a_{lm}^{(k)}$ is the (l,m) element of the matrix A^k. Define $a_{11}^{(0)} = 1$ and $a_{21}^{(0)} = 0$ and the transformation

$$a_{bk} = \frac{1+i}{\phi_\pi} a_{11}^{(k)}.$$

The sequence $\{a_{bk}\}_{k=0}^{\infty}$ defines the Impulse Response Functions of $\hat{b}_{ht}$.

First, I will show that $a_{bk} \geq 0$ for all $k = 0, 1, 2, \ldots$ when the matrix A is diagonalizable, and then I will handle the case when the eigenvalue is repeated and A is not diagonalizable (the only other case we need to worry about for a two by two matrix).

Assuming that A is diagonalizable, define

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

as a matrix with the two eigenvalues of A on its diagonal ordered like $\lambda_1 > \lambda_2$ (remember we are handling the case of real eigenvalues right now) and P as a matrix that has the eigenvectors of A as its columns. Since we have assumed A is diagonalizable, we have $A = P\Lambda P^{-1}$ and also $A^k = P\Lambda^k P^{-1}$. Since Λ is diagonal

$$\Lambda^k = \begin{bmatrix} \lambda_1^k & 0 \\ 0 & \lambda_2^k \end{bmatrix}$$

and thus if we expand the expression for A^k we obtain that

$$a_{11}^{(k)} = \frac{p_{11}p_{22}\lambda_1^k - p_{12}p_{21}\lambda_2^k}{|P|}$$

where $|P|$ is the determinant of the matrix of eigenvectors P and p_{lm} is its (l,m)-th element. Since both of the eigenvalues are positive and are ordered so that $\lambda_1 > \lambda_2$ it follows that $|P| > 0$ and hence

$$\frac{p_{11}p_{22}\lambda_1^k - p_{12}p_{21}\lambda_2^k}{|P|} > 0.$$

This proves that $a_{11}^{(k)} > 0$ for all k and hence $a_{bk} > 0$ for all k. This completes the proof for diagonalizable A – now assume that A is not diagonalizable. We can instead use the Jordan Decomposition to again write $A = P\Lambda P^{-1}$ but now

$$\Lambda = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$$

and the columns of P are the generalized eigenvectors of A. In this case, there is only one linearly independent eigenvector associated with the eigenvalue of λ, call it $\vec{p}$, and thus the second generalized eigenvector, call it $\vec{u}$, is a 2x1 vector that solves

$$(A - \lambda I)\vec{u} = \vec{p}$$

We can solve for the needed eigenvectors via standard techniques, and obtain $\vec{p} = [p_1, 1]'$ and $\vec{u} = [u_1, 1]'$, where $p_1 = \frac{\lambda - \rho_\tau}{(1-\rho_\tau)\kappa_b \frac{\tau}{b_h}}$, $u_1 = p_1 + \frac{1}{(1-\rho_\tau)\kappa_b \frac{\tau}{b_h}}$. We can then use $A^k = P\Lambda^k P^{-1}$ to get:

$$a_{11}^{(k)} = \lambda^{k-1}(\lambda + k\frac{p_1}{u_1 - p_1}) > 0$$

The inequality follows from $u_1 > p_1 > 0$, $\lambda > 0$. This completes the proof of part (i).

Part (ii): From the proof of Lemma 1 we know that $\rho_\tau \in (\underline{\rho}_\tau, \frac{\theta_2}{\theta})$ implies that the eigenvalues of A are complex. We can express them as $\lambda_1 = a + bi$ and $\lambda_2 = a - bi$ where $a = \frac{1}{2}(\theta - (1-\rho_\tau)\kappa_b + \rho_\tau\theta_2) > 0$, $b = \frac{1}{2}\sqrt{4\theta\theta_2\rho_\tau - (\theta - (1-\rho_\tau)\kappa_b + \rho_\tau\theta_2)^2} > 0$ and i is the imaginary unit. The two conjugate eigenvectors can be written as $\vec{p}_k = [x \pm yi, 1]'$, where .

$$x = \frac{\tau}{b_h}\frac{(\theta - (1-\rho_\tau)\kappa_b + \rho_\tau\theta_2 - 2\rho_\tau)}{2(1-\rho_\tau)\kappa_b}$$

$$y = \frac{\tau}{b_h}\frac{\sqrt{4\theta\theta_2\rho_\tau - (\theta - (1-\rho_\tau)\kappa_b + \rho_\tau\theta_2)^2}}{2b(1-\rho_\tau)\kappa_b}$$

With two conjugate complex eigenvalues A is diagonalizable and can be expressed as $A = P\Lambda P^{-1}$ where P is a similarity matrix with the eigenvectors of A as its columns and Λ is a diagonal matrix with the eigenvalues on the diagonal. By Euler's formula $\lambda_1 = a + bi = |\lambda|e^{\zeta i}$ where $\zeta = \arctan(\frac{b}{a})$ and $|\lambda| = \sqrt{a^2 + b^2}$ is the magnitude of the complex roots. This formulation is convenient because it is easy to take powers of the eigenvalues, (e.g. $\lambda_1^k = |\lambda|^k e^{k\zeta i}$) and hence it is easy to compute powers of the eigenvalue matrix Λ. Using this, Euler's formula and the fact that $A^k = P\Lambda^k P^{-1}$ it is straightforward to compute

$$\begin{aligned}
a_{11}^{(k)} &= |\lambda|^k(\cos(k\zeta) + \frac{x}{y}\sin(k\zeta)) \\
&= |\lambda|^k\sqrt{1 + (\frac{x}{y})^2}\sin(k\zeta + \psi) \\
&= |\lambda|^k\sqrt{1 + (\frac{x}{y})^2}\cos(k\zeta + \psi - \frac{\pi}{2})
\end{aligned}$$

where $\psi = \arctan(\frac{y}{x}) + \pi\mathbb{I}(\frac{y}{x} < 0)$. The second equality follows from the formula for linear combinations of trig functions, and the third is simply an application of $\cos(\theta - \frac{\pi}{2}) = \sin(\theta)$.

By the definition of the $\arctan(\cdot)$ function and the virtue of $a \geq 0, b \geq 0$ it follows that $\zeta \in [0, \frac{\pi}{2})$. If $\kappa_b \leq \frac{\theta + (\theta_2 - 2)\rho_\tau}{1-\rho_\tau}$, then $x \geq 0$ and $\psi \leq \frac{\pi}{2}$ and this case $\cos(k\zeta + \psi - \frac{\pi}{2}) \geq 0$ for at least $k = 1$. Otherwise, use the formula for addition of arctangent to get,

$$\arctan(\frac{b}{a}) + \arctan(\frac{y}{x}) = \arctan(\frac{\frac{b}{a} + \frac{y}{x}}{1 - \frac{by}{ax}}).$$

where $1 - \frac{by}{ax} > 0$. And since $\kappa_b \in (\frac{\theta + (\theta_2 - 2)\rho_\tau}{1-\rho_\tau}, \frac{\theta + (\theta_2 - 1)\rho_\tau}{1-\rho_\tau})$, we can show that $\frac{b}{a} + \frac{y}{x} < 0$ and therefore $\arctan(\frac{\frac{b}{a} + \frac{y}{x}}{1 - \frac{by}{ax}}) \in (-\frac{\pi}{2}, 0)$. Therefore, we again reach the conclusion that $\cos(k\zeta + \psi - \frac{\pi}{2}) \geq 0$ for at least $k = 1$. This completes the proof of Lemma 2.

Lastly, if $\phi_\pi < 1, \kappa_b \in [0, \theta - \theta_2), \rho_\tau \in [0, 1)$, then from the proof of Lemma 1 we know that $\kappa_b < \theta - \theta_2$ ensures the eigenvalues are real, and as we saw from the proof of Lemma 2, in this case the IRF never crosses the steady state. In fact, from the proof of Lemma 1 we also have the stronger result that $\hat{b}_{ht} = 0$, and hence the IRF is

$$a_{bk} = 0 \text{ for } k = 0, 1, 2, 3, \ldots$$

□

C.3 PROPOSITION 1:

PROPOSITION 1. *The sign of the UIP regression coefficients* $\beta_k = \frac{\text{Cov}(\hat{\lambda}_{t+k}, \hat{\imath}_t - \hat{\imath}_t^*)}{\text{Var}(\hat{\imath}_t - \hat{\imath}_t^*}$ *depend on the monetary-fiscal policy mix as follows.*

(i) ***Active Monetary, Passive Fiscal policy*** $(\phi_\pi > 1, \kappa_b \in (\theta - \theta_2, \frac{\theta + (\theta_2 - 1)\rho_\tau}{1 - \rho_\tau}))$**:**

  (a) $\rho_\tau \leq \underline{\rho}$*: classic UIP puzzle at all horizons:* $\beta_k < 0$ *for* $k = 1, 2, 3, \ldots$

  (b) $\rho_\tau > \underline{\rho}$*: UIP violations exhibit cyclical dynamics and eventually switch sign:*

$$\beta_k < 0 \textit{ for } k < \bar{k} \textit{ and } \beta_k > 0 \textit{ for some } k > \bar{k}$$

(ii) ***Passive Monetary, Active Fiscal policy*** $(\phi_\pi < 1, \kappa_b \in (0, \theta - \theta_2))$**:** *UIP violations go in the same direction at all horizons and are in fact always zero:*

$$\beta_k = 0 \textit{ for } k = 1, 2, 3, \ldots$$

Proof. **Part (i), sub-point (a):** Start with the definition of the UIP regression coefficient,

$$\beta_k = \frac{\text{Cov}(\hat{\lambda}_{t+k}, \hat{\imath}_t - \hat{\imath}_t^*)}{\text{Var}(\hat{\imath}_t - \hat{\imath}_t^*)}$$

and note that in equilibrium the expected excess returns are linear in bond holdings,

$$E_t(\hat{\lambda}_{t+1}) = -\chi_b \hat{b}_{ht} \tag{8}$$

where $\chi_b = -\frac{b_h}{\beta(1+i)u_c}\left((u_{b_h b_h} - \frac{u_{b_h} u_{cb_h}}{u_c}) - (u_{b_h b_f} - \frac{u_{b_f} u_{cb_h}}{u_c})\right) s$. Where $u_x(.)$ and $u_{xy}(.)$ respectively are the steady state values of the first and second partial derivative of the utility function. In the symmetric steady state, $u_{b_f} = u_{b_h}$ and given the assumption of imperfect substitutability between home and foreign bonds (and since utility is concave):

$$u_{b_h b_h} < u_{b_f b_f} < 0$$

it follows that $\chi_b > 0$. By Lemma 2, we know that in this case (Active Monetary policy), the IRF of $\hat{b}_{ht}$ is positive at all horizons (i.e. $a_{bk} > 0$ for all k), and next, I will show that the IRF of the interest rate differential $\hat{i}_t - \hat{i}_t^*$ is also always positive. Then by (8) we can conclude that $\beta_k < 0$ for all $k \geq 1$.

To derive the IRF of the interest rate differential, note that since the foreign interest rate is constant, $\hat{i}_t - \hat{i}_t^* = \hat{i}_t = \phi_\pi + v_t$. From Lemma 1, the equilibrium inflation is given by

$$\hat{\pi}_t = \gamma_b^\pi \hat{b}_{ht} + \gamma_\tau^\pi \hat{\tau}_t - \frac{v_t}{\phi_\pi}$$

where $\gamma_b^\pi = \gamma_M + \frac{\theta_2(\phi_\pi - \rho_\tau)(\gamma_\Psi - \gamma_M(\phi_\pi - 1))}{\phi_\pi(\kappa_b(1-\rho_\tau)+\theta_2(\phi_\pi-\rho_\tau))-\theta(\phi_\pi-\rho_\tau)} > 0$, $\gamma_\tau^\pi = -\frac{\tau}{b_h}\frac{\rho_\tau(\gamma_\Psi - \gamma_M(\phi_\pi-1))}{\phi_\pi(\kappa_b(1-\rho_\tau)+\theta_2(\phi_\pi-\rho_\tau))-\theta(\phi_\pi-\rho_\tau)}$. Thus,

$$\begin{aligned}
\hat{i}_t - \hat{i}_t^* &= \phi_\pi(\gamma_b^\pi \hat{b}_{ht} + \gamma_\tau^\pi \hat{\tau}_t) \\
&= \phi_\pi((\gamma_b^\pi a_{b0} + \gamma_\tau^\pi a_{\tau 0})v_t + (\gamma_b^\pi a_{b1} + \gamma_\tau^\pi a_{\tau 1})v_{t-1} + \dots) \\
&= a_{i0}v_t + a_{i1}v_{t-1} + a_{i2}v_{t-2} + \dots
\end{aligned}$$

where I have substituted in the Wold decomposition of $\hat{b}_{ht}$ and $\hat{\tau}_t$, and by the proof of Lemma 2, $a_{bk} = \frac{1+i}{\phi_\pi}a_{11}^{(k)}$ and $a_{\tau k} = \frac{1+i}{\phi_\pi}a_{21}^{(k)}$, with $a_{lm}^{(k)}$ the (k,l) element of the matrix A^k. This defines the Wold decomposition of the interest rate differential through the coefficients a_{ik}, where

$$\begin{aligned}
a_{ik} = \phi_\pi(\gamma_b^\pi a_{bk} + \gamma_\tau^\pi a_{\tau k}) &= (1+i)\left(\gamma_b^\pi \frac{\lambda_1^k p_{11} - \lambda_2^k p_{22}}{|P|} + \gamma_\tau^\pi \frac{\lambda_1^k - \lambda_2^k}{|P|}\right) \\
&= (1+i)\left(\frac{\lambda_1^k}{|P|}(p_{11}\gamma_b^\pi + \gamma_\tau^\pi) - \frac{\lambda_2^k}{|P|}(p_{12}\gamma_b^\pi + \gamma_\tau^\pi)\right)
\end{aligned}$$

and $\lambda_1 > \lambda_2 > 0$ are the ordered eigenvalues of A, and P is the matrix of eigenvectors, with $p_{11} = \frac{\lambda_1 - \rho_\tau}{(1-\rho_\tau)\kappa_b}\frac{b_h}{\tau}$, and $p_{12} = \frac{\lambda_2 - \rho_\tau}{(1-\rho_\tau)\kappa_b}\frac{b_h}{\tau}$. Since the eigenvalues are ordered and positive, $p_{11} > p_{12} > 0$, and hence $|p_{11}\gamma_b^\pi + \gamma_\tau^\pi| > |p_{12}\gamma_b^\pi + \gamma_\tau^\pi|$. If $\gamma_\tau^\pi > 0$ then it follows that $p_{11}\gamma_b^\pi + \gamma_\tau^\pi > 0$, and thus $\left(\frac{\lambda_1^k}{|P|}(p_{11}\gamma_b^\pi + \gamma_\tau^\pi) - \frac{\lambda_2^k}{|P|}(p_{12}\gamma_b^\pi + \gamma_\tau^\pi)\right) > 0$ and hence $a_{ik} > 0$.

On the other hand, if $\gamma_\tau^\pi < 0$, first we need to show $p_{11}\gamma_b^\pi + \gamma_\tau^\pi > 0$. Start with,

$$\begin{aligned}
p_{11}\gamma_b^\pi - |\gamma_\tau^\pi| \propto\ & (\theta - \kappa_b(1-\rho_\tau) - \theta_2\rho_\tau + \sqrt{(\theta - \kappa_b(1-\rho_\tau) + \theta_2\rho_\tau)^2 - 4\theta\theta_2\rho_\tau})(\gamma_\Psi(\phi_\pi - \rho_\tau) - \gamma_M(i(\phi_\pi - \rho_\tau) - \kappa_b(1-\rho_\tau)\phi_\pi)) \\
& - 2\kappa_b\theta_2(1-\rho_\tau)\rho_\tau(\gamma_\Psi + \gamma_M(1-\phi_\pi)) \\
\geq\ & (\theta - \kappa_b(1-\rho_\tau) - \theta_2\rho_\tau)(\gamma_\Psi(\phi_\pi - \rho_\tau) - \gamma_M(i(\phi_\pi - \rho_\tau) - \kappa_b(1-\rho_\tau)\phi_\pi)) - 2\kappa_b\theta_2(1-\rho_\tau)\rho_\tau(\gamma_\Psi + \gamma_M(1-\phi_\pi))
\end{aligned}$$

The last equation is concave and quadratic in κ_b, so if it is positive for any $k_1 < k_2$, then it's positive for all values in between as well. Furthermore, note that in order for the the eigenvalues to be real and less than one in magnitude we must have $\kappa_b \in (\theta - \theta_2, \frac{\theta + \theta_2\rho_\tau - 2\sqrt{\theta\theta_2\rho_\tau}}{1-\rho_\tau}]$, and thus it is enough to show that the quadratic equation is positive at both ends of this interval.

For $\kappa_b = \theta - \theta_2$,

$$p_{11}\gamma_b^\pi - |\gamma_\tau^\pi| \geq \gamma_\Psi(1-\rho_\tau)(\theta_2 + \theta\rho_\tau - 2\theta_2(1+\theta-\theta_2)\rho_\tau) + \gamma_M((\theta - \theta_2 - i)(1-\rho_\tau)(\theta_2 + \theta\rho_\tau - 2\theta_2\rho_\tau)$$

and since $\rho_\tau \in [0, \frac{\theta_2}{\theta})$ it follows that $(\theta_2 + \theta\rho_\tau - 2\theta_2\rho_\tau) > 0$, and $(\theta_2 + \theta\rho_\tau - 2\theta_2(1+\theta-\theta_2)\rho_\tau) > 0$. Also $\theta - \theta_2 - i = (1+i)\gamma_\Psi > 0$, and hence $p_{11}\gamma_b^\pi - |\gamma_\tau^\pi| > 0$.

On the other hand, if $\kappa_b = \frac{\theta + \theta_2\rho_\tau - 2\sqrt{\theta\theta_2\rho_\tau}}{1-\rho_\tau}$:

$$\begin{aligned}
p_{11}\gamma_b^\pi - |\gamma_\tau^\pi| &\geq 2\gamma_\Psi((1-\rho_\tau)\sqrt{\theta\theta_2\rho_\tau} - \theta_2\rho_\tau(\theta_2\rho_\tau + 1 + \theta - \rho_\tau - 2\sqrt{\theta\theta_2\rho_\tau}) + 2\gamma_M(\sqrt{\theta\theta_2\rho_\tau} - \theta_2\rho_\tau)(\theta + \theta_2\rho_\tau - 2\sqrt{\theta\theta_2\rho_\tau} - (1-\rho_\tau)i) \\
&= 2\gamma_\Psi((1-\rho_\tau)(\sqrt{\theta\theta_2\rho_\tau} - \theta_2\rho_\tau) - \theta_2\rho_\tau(\theta - 2\sqrt{\theta\theta_2\rho_\tau} + \theta_2\rho_\tau)) + 2\gamma_M\sqrt{\theta_2\rho_\tau}(\sqrt{\theta} - \sqrt{\theta_2\rho_\tau})(\theta + \theta_2\rho_\tau - 2\sqrt{\theta\theta_2\rho_\tau} - (1-\rho_\tau)i) \\
&= 2\sqrt{\theta_2\rho_\tau}(\sqrt{\theta} - \sqrt{\theta_2\rho_\tau})\Big(\gamma_\Psi((1-\rho_\tau) - \sqrt{\theta_2\rho_\tau}(\sqrt{\theta} - \sqrt{\theta_2\rho_\tau})) + \gamma_M((\sqrt{\theta} - \sqrt{\theta_2\rho_\tau})^2 - i(1-\rho_\tau))\Big) \\
&= 2\sqrt{\theta_2\rho_\tau}(\sqrt{\theta} - \sqrt{\theta_2\rho_\tau})\underbrace{\Big((\gamma_\Psi - \gamma_M i)(1-\rho_\tau) + \gamma_\Psi\theta_2\rho_\tau + \gamma_M(\theta + \theta_2\rho_\tau) - (2\gamma_M + \gamma_\Psi)\sqrt{\theta\theta_2\rho_\tau}\Big)}_{=\Omega}
\end{aligned}$$

Since $\rho_\tau < \frac{\theta_2}{\theta}$ and $\theta_2 < \theta_1$ it follows that $\sqrt{\theta} > \sqrt{\theta_2\rho_\tau}$. To evaluate the second piece in parenthesis (which I have named Ω for brevity), substitute in $\theta = (1+i)(1+\gamma_\Psi+\gamma_M)$ and $\theta_2 = 1 + \gamma_M(1+i)$ and simplify to get:

$$\Omega = ((\gamma_\Psi + \gamma_M)(1 + \gamma_M(1+i)) + \gamma_M(1+\gamma_\Psi+\gamma_M)(1+i)\rho_\tau)^2 - (1+\gamma_\Psi+\gamma_M)(\gamma_\Psi+2\gamma_M)^2(1+i)(1+\gamma_M(1+i))\rho_\tau$$

This is a convex quadratic equation in ρ_τ, with zeros at $\rho_1 = \frac{\theta_2}{\theta}$ and at $\rho_2 = \frac{\theta_2}{\theta}\frac{(\gamma_\Psi+\gamma_M)^2}{\gamma_M^2}$, and since $\gamma_\Psi > 0$, $\rho_1 < \rho_2$. Therefore, $\Omega > 0$ for all $\rho_\tau < \frac{\theta_2}{\theta}$ and thus we conclude that $p_{11}\gamma_b^\pi - |\gamma_\tau^\pi| > 0$.

Thus, we have shown that under Active Monetary Policy we have $p_{11}\gamma_b^\pi - |\gamma_\tau^\pi| > 0$, and thus since $\lambda_1 > \lambda_2 > 0$ we have $\left(\frac{\lambda_1^k}{|P|}(p_{11}\gamma_b^\pi + \gamma_\tau^\pi) - \frac{\lambda_2^k}{|P|}(p_{12}\gamma_b^\pi + \gamma_\tau^\pi)\right) > 0$. Therefore, under Active Monetary policy, $a_{ik} > 0$ for all k.

Plugging this and the IRF for $\hat{b}_{ht}$ in the UIP regression coefficients, I obtain

$$\begin{aligned}
\beta_k &= -\chi_b\frac{\mathrm{Cov}(\hat{b}_{h,t+k-1}, \hat{i}_t - \hat{i}_t^*)}{\mathrm{Var}(\hat{i}_t - \hat{i}_t^*)} \\
&= -\chi_b\frac{\sigma_v^2(a_{b,k-1}a_{i,0} + a_{b,k}a_{i,1} + a_{b,k+1}a_{i,2} + \dots)}{\mathrm{Var}(\hat{i}_t - \hat{i}_t^*)} \\
&< 0
\end{aligned}$$

where the inequality follows from $\chi_b > 0$ and $a_{bk} > 0$ and $a_{ik} > 0$ for all k.

Above we implicitly assumed that A is diagonalizable. But the proof is very similar if it is not, with the only difference being that

$$a_{ik} = \phi_\pi(\gamma_b^\pi a_{bk} + \gamma_\tau^\pi a_{\tau k}) = (1+i)\left(\gamma_b^\pi \lambda^{k-1}(\lambda + k\frac{p_{11}}{p_{12}-p_{11}}) + \gamma_\tau^\pi k \frac{\lambda^{k-1}}{p_{12}-p_{11}}\right)$$
$$= (1+i)\left(\gamma_b^\pi \lambda^k + \frac{k\lambda^{k-1}}{p_{12}-p_{11}}(\gamma_b^\pi p_{11} + \gamma_\tau^\pi)\right)$$

But we have already shown $(\gamma_b^\pi p_{11} + \gamma_\tau^\pi) > 0$, and by the proof of Lemma 2, $p_{12} - p_{11} > 0$, hence $a_{ik} > 0$ for all k again, and we are done.

Part (i), sub-point (b): Here I work under the assumption that the roots are complex - i.e. $\rho_\tau > \underline{\rho}(\kappa_b)$ as defined in Lemma 2. We can express the UIP regression coefficients as

$$\beta_k = \frac{\text{Cov}(-\chi_b E_t(\hat{b}_{h,t+k-1}), \phi_\pi(\gamma_b^\pi \hat{b}_{ht} + \gamma_\tau^\pi \hat{\tau}_t))}{\text{Var}(\phi_\pi(\gamma_b^\pi \hat{b}_{ht} + \gamma_\tau^\pi \hat{\tau}_t))} = -\chi_b \phi_\pi(\gamma_b^\pi \frac{\text{Cov}(E_t(\hat{b}_{h,t+k-1}), \hat{b}_{ht})}{\text{Var}(\phi_\pi(\gamma_b^\pi \hat{b}_{ht} + \gamma_\tau^\pi \hat{\tau}_t))} + \gamma_\tau^\pi \frac{\text{Cov}(E_t(\hat{b}_{h,t+k-1}), \hat{\tau}_{ht})}{\text{Var}(\phi_\pi(\gamma_b^\pi \hat{b}_{ht} + \gamma_\tau^\pi \hat{\tau}_t))})$$

Since $E_t(\hat{b}_{t+k}) = [1,0]A^k x_t$, we have

$$\text{Cov}(E_t(\hat{b}_{t+k}), b_t) = a_{11}^{(k)} \text{Var}(\hat{b}_t) + a_{12}^{(k)} \text{Cov}(\hat{b}_t, \hat{\tau}_t) \tag{9}$$
$$\text{Cov}(E_t(\hat{b}_{t+k}), \tau_t) = a_{11}^{(k)} \text{Cov}(\hat{b}_t, \hat{\tau}_t) + a_{12}^{(k)} \text{Var}(\hat{\tau}_t) \tag{10}$$

Compute the variance on both sides of the tax policy rule to obtain

$$Var(\hat{\tau}_t) = \frac{b_h^2}{\tau^2}\frac{k_b^2}{(1+\rho_\tau)}(1-\rho_\tau)Var(\hat{b}_t) + 2\frac{b_h}{\tau}\frac{\kappa_b\rho_\tau}{(1+\rho_\tau)}Cov(\hat{\tau}_t, b_t)$$

and then combine with

$$\text{Cov}(\hat{\tau}_t, \hat{b}_t) = \text{Cov}(\rho\hat{\tau}_{t-1} + a_{21}^{(1)}\hat{b}_{t-1}, a_{11}^{(1)}\hat{b}_{t-1} + a_{12}^{(1)}\hat{\tau}_{t-1} + \frac{1+i}{\phi_\pi}v_t)$$
$$= -\frac{\rho_\tau^2}{\theta_2}\frac{\tau}{b}\text{Var}(\hat{\tau}_t) + \frac{\theta-(1-\rho_\tau)\kappa_b}{\theta_2}(1-\rho_\tau)\kappa_b\frac{b}{\tau}\text{Var}(\hat{b}_t) + (\frac{\theta-(1-\rho_\tau)\kappa_b}{\theta_2}\rho_\tau - (1-\rho_\tau)\kappa_b\frac{\rho_\tau}{\theta_2})\text{Cov}(\hat{b}_t, \hat{\tau}_t)$$

to obtain

$$\text{Cov}(\hat{\tau}_t, \hat{b}_t) = \underbrace{(1-\rho_\tau)\kappa_b\frac{b_h}{\tau}\frac{(\theta(1+\rho_\tau)-\kappa_b)}{\theta_2 + \rho_\tau(\theta_2 + 2\kappa_b - \theta(1+\rho_\tau))}}_{=\delta}\text{Var}(\hat{b}_t).$$

Substituting this back in (9) yields $\text{Cov}(E_t(\hat{b}_{t+k}), b_t) = (a_{11}^{(k)} + \delta a_{12}^{(k)})\,\text{Var}(\hat{b}_{ht})$, and similarly substituting things out in (10) yields $\text{Cov}(E_t(\hat{b}_{t+k}), \hat{\tau}_t) = (a_{11}^{(k)}\delta + a_{12}^{(k)}((\frac{b_h}{\tau})^2 \frac{\kappa_b^2(1-\rho_\tau)}{1+\rho_\tau} + 2\frac{b_h}{\tau}\frac{\kappa_b\rho_\tau\delta}{1+\rho_\tau}))\,\text{Var}(\hat{b}_{ht})$, and hence the UIP coefficient becomes

$$
\begin{aligned}
\beta_{k+1} &= -\frac{\chi_b\phi_\pi \,\text{Var}(\hat{b}_{ht})}{Var(\hat{i}_t)}(\gamma_b^\pi(a_{11}^{(k)} + \delta a_{12}^{(k)}) + \gamma_\tau^\pi(a_{11}^{(k)}\delta + a_{12}^{(k)}((\frac{b_h}{\tau})^2\frac{\kappa_b^2(1-\rho_\tau)}{1+\rho_\tau} + 2\frac{b_h}{\tau}\frac{\kappa_b\rho_\tau\delta}{1+\rho_\tau})))\\
&= -\frac{\chi_b\phi_\pi \,\text{Var}(\hat{b}_{ht})}{Var(\hat{i}_t)}\left(a_{11}^{(k)}\underbrace{(\gamma_b^\pi + \gamma_\tau^\pi\delta)}_{=\gamma_{a_{11}}} + a_{12}^{(k)}\underbrace{(\gamma_b^\pi\delta + \gamma_\tau^\pi(((\frac{b_h}{\tau})^2\frac{\kappa_b^2(1-\rho_\tau)}{1+\rho_\tau} + 2\frac{b_h}{\tau}\frac{\kappa_b\rho_\tau\delta}{1+\rho_\tau})))}_{=\gamma_{a_{12}}}\right)\\
&= -\frac{\chi_b\phi_\pi \,\text{Var}(\hat{b}_{ht})}{Var(\hat{i}_t)}\left(a_{11}^{(k)}\gamma_{a_{11}} + a_{12}^{(k)}\gamma_{a_{12}}\right)\\
&= -\frac{\chi_b\phi_\pi \,\text{Var}(\hat{b}_{ht})}{Var(\hat{i}_t)}\left(\gamma_{a_{11}}|\lambda|^k(\cos(k\zeta) + \frac{x}{y}\sin(k\zeta)) - \gamma_{a_{12}}|\lambda|^k\frac{x^2+y^2}{y^2}\sin(k\zeta)\right)\\
&= -\frac{\chi_b\phi_\pi \,\text{Var}(\hat{b}_{ht})}{Var(\hat{i}_t)}|\lambda|^k\left(\gamma_{a_{11}}\cos(k\zeta) + (\gamma_{a_{11}}\frac{x}{y} - \gamma_{a_{12}}\frac{x^2+y^2}{y^2})\sin(k\zeta))\right)\\
&= -\frac{\chi_b\phi_\pi \,\text{Var}(\hat{b}_{ht})}{Var(\hat{i}_t)}|\lambda|^k\underbrace{\sqrt{\gamma_{a_{11}}^2 + (\gamma_{a_{11}}\frac{x}{y} - \gamma_{a_{12}}\frac{x^2+y^2}{y^2})^2}}_{=\Gamma}\cos(k\zeta + \psi - \frac{\pi}{2})\\
&= -\frac{\chi_b\phi_\pi \,\text{Var}(\hat{b}_{ht})}{Var(\hat{i}_t)}|\lambda|^k\Gamma\cos(k\zeta + \psi - \frac{\pi}{2})
\end{aligned}
$$

where $\psi = \arctan(\frac{\gamma_{a_{11}}}{\gamma_{a_{11}}\frac{x}{y} - \gamma_{a_{12}}\frac{x^2+y^2}{y^2}}) + \pi\mathbb{I}(\frac{\gamma_{a_{11}}}{\gamma_{a_{11}}\frac{x}{y} - \gamma_{a_{12}}\frac{x^2+y^2}{y^2}} < 0)$, and x and y are the real and imaginary part of the eigenvectors as defined in Lemma 2, and $\zeta = \arctan(\frac{b}{a}) \in [0, \frac{\pi}{2})$ where the eigenvalue is $a + bi$. I am also using the convention that $a_{11}^{(0)} = 1$ and $a_{12}^{(0)} = 0$.

This gives us the general expression of β_k and shows that it is cyclical, and changes sign as the cosine expression changes sign. Lastly, I will show that $\beta_1 < 0$, which finishes the proof by establishing that the regression coefficients start negative, and then will eventually turn positive as k grows (since $\zeta \in [0, \frac{\pi}{2})$).

To show $\beta_1 < 0$, start by re-writing it as $\beta_1 = -\frac{\chi_b\phi_\pi \,\text{Var}(\hat{b}_{ht})}{Var(\hat{i}_t)}(\gamma_b^\pi + \gamma_\tau^\pi\delta)$ by using the fact that $a_{11}^{(0)} = 1$ and $a_{12}^{(0)} = 0$, and notice that it is enough to show that $\gamma_b^\pi + \delta\gamma_\tau^\pi > 0$. Substitute in the definitions for the three variables, bring everything to a common denominator, and since the resulting denominator is positive, the sign of $\gamma_b^\pi + \delta\gamma_\tau^\pi$ is the same as the sign of the numerator:

$$
(\theta_2(1+\rho_\tau) + \rho_\tau(2\kappa_b - \theta(1+\rho_\tau))(\gamma_\Psi(\phi_\pi - \rho_\tau) + \gamma_M\kappa_b(1-\rho_\tau) - \gamma_M(\phi_\pi - \rho_\tau)) - \kappa_b(1-\rho_\tau)\rho_\tau(\theta(1+\rho_\tau - \kappa_b)(\gamma_\Psi - \gamma_M(\phi_\pi - 1)) \tag{11}
$$

This is a convex quadratic function of κ_b ($\frac{\partial^2 .}{(\partial \kappa_b)^2} = 2(1-\rho_\tau)\rho_\tau(\gamma_\Psi + \gamma_M(\phi_\pi + 1)) > 0$), and I will show that it is positive for all $\kappa_b > \theta - \theta_2$, by showing that it is positive and increasing at $\kappa_b = \theta - \theta_2$.

At $\kappa_b = \theta - \theta_2$, the expression becomes

$$\gamma_\Psi(1-\rho_\tau)(\theta_2 + \theta\rho_\tau)(\phi_\pi\theta_2 - \rho_\tau\theta) > \gamma_\Psi(1-\rho_\tau)(\theta_2 + \theta\rho_\tau)(\phi_\pi\theta_2 - \theta_2) > 0$$

where the first inequality follows from $\rho_\tau < \frac{\theta_2}{\theta}$, and the second from $\phi_\pi > 1$.

On the other hand, its derivative at $\kappa_b = \theta - \theta_2$ is:

$$\begin{aligned}\gamma_\Psi\rho_\tau(\theta(1-\rho_\tau)^2 + 2(\phi_\pi - \rho_\tau - \theta_2(1-\rho_\tau))) + \gamma_M(\theta(1-\rho_\tau)\rho_\tau(2\phi_\pi + 1 - \rho_\tau) + 2i\rho_\tau(\phi_\pi - \rho_\tau) + \theta_2(1-\rho_\tau)(\phi_\pi(1-\rho_\tau) - 2\rho_\tau))\\ > (1-\rho_\tau)^2(\gamma_M\theta_2 + (\gamma_\Psi + \gamma_M)\theta\rho_\tau)\\ > 0\end{aligned}$$

where the first inequality follows from the fact that the top line is increasing in ϕ_π and $\phi_\pi > 1$. Thus, we have shown that (11) is positive and increasing at $\kappa_b = \theta - \theta_2$, and hence $\gamma_b^\pi + \delta\gamma_\tau^\pi > 0$ which implies that $\beta_1 > 0$. This completes the proof of part (i), sub-point b.

Part (ii): By the last part of the proof of Lemma 2, the eigenvalues of A are always real in this case, and by similar steps to the proof of Proposition 1, Part (i), sub-point (a) we can show that the IRF of $\hat{i}_t$ is positive at all horizons and hence β_k has the same sign for all k. Moreover, from Lemma 3 we have the particular result that $\hat{b}_{h,t+k} = 0$ for all k, and hence

$$\beta_k = -\chi_b \frac{\text{Cov}(\hat{b}_{h,t+k-1}, \hat{i}_t - \hat{i}_t^*)}{\text{Var}(\hat{i}_t - \hat{i}_t^*)} = -\chi_b \frac{\text{Cov}(0, \hat{i}_t - \hat{i}_t^*)}{\text{Var}(\hat{i}_t - \hat{i}_t^*)} = 0$$

□

D Model Discussion

D.1 Forward Exchange Rate Contracts and UIP Violations

In this section, I augment the model to include trade in forward contracts on currencies, and show that the convenience yield mechanism generates UIP violations that emerge both when looking at exchange rates and interest rates data only, and when looking at forward and spot exchange rates.

With trade in forward markets, the compensating financial return that arises from non-zero convenience yield differentials can come from both UIP and CIP deviations. When a US investors buys a foreign bond (say Euro bond) and sells the proceeds forward at rate F_t, he creates a synthetic USD safe asset, because he now has safe stream of future USD payments. This is similar to a US Treasury, except for the fact that it is a synthetic position, as opposed to holding the physical US Treasury.

Let Ψ_t^{US} be the convenience yield on the Treasury and $\tilde{\Psi}_t^{US}$ the convenience yield on the synthetic US safe bond, constructed with a long position in a Euro bond whose currency risk has been sold forward. No arbitrage in forward markets requires that

$$i_t^{US} + \Psi_t^{US} = f_t - s_t + i_t^{EUR} + \tilde{\Psi}_t^{US}$$

If the synthetic position has a similar convenience yield to the physical bond $\Psi_t^{US} \approx \tilde{\Psi}_t^{US}$ then CIP will approximately hold, otherwise there will be a CIP deviation.

Since empirically CIP has been shown to hold very well for the entirety of my sample – it has only exhibited non-trivial deviations since the financial crisis (see Du et al. (2018)), which is not part of the main sample – Iit appears that in normal times (pre-2008) the synthetic position indeed possesses the same convenience yield as owning the physical bond. The opening up of a persistent CIP deviation post–2008, and some preliminary evidence that carry trade profits have declined since the crisis, could be evidence that markets have since switched to a different equilibrium where the convenience yield drives a wedge in the CIP condition. This could be an interesting avenue for future research but is outside of the scope of the current paper, which will focus on the equilibrium where CIP holds.

Thus, when the synthetic and physical assets have the same convenience yield we have the familiar result that

$$f_t - s_t = i_t^{US} - i_t^{EUR}$$

This leads to the important result that (in log-approximation) the expected return on buying foreign currency forward (a popular way of implementing the carry trade without the need to transact in bond markets) is simply the exchange rate change plus the interest rate differential on the underlying assets. Thus, taking a position in forwards is equivalent to buying and selling the physical assets, hence:

$$E_t(s_{t+1} - f_t) = E_t(\Delta s_{t+1} + i_t^* - i_t) = \hat{\Psi}_{H,t} - \hat{\Psi}_{F,t}.$$

At the end of the day, the strategy implemented through forwards market has equivalent financial and convenience returns to a trade in the home and foreign bonds themselves, hence the forwards data would display equivalent UIP violations and the mechanism works in the same way. Due to this equivalence and for simplicity, the benchmark model abstracts from trade in forward contracts.

D.2 Shutting down convenience yields

The convenience yield mechanism helps the model fit the joint behavior of interest rates and exchange rates, but otherwise has a small effect on the implications of other macro aggregates. In Table 1 I report the resulting moments of a model where the convenience yield mechanism is shut down completely, side-by-side with the moments of the benchmark calibration and the data.

Interestingly, shutting down the convenience yield mechanism has only a relatively small effect on most moments. Clearly, this model has no excess currency returns or UIP deviations. However, there is relatively little change in the results for macro aggregates like output, consumption and government debt. In fact, the main change is in the lower persistent and volatility of interest rates and interest rate differentials (both nominal and real).

The intuition for the relatively minor changes has to do with the fact that the long-run behavior of the real exchange rate is roughly consistent with UIP, because the negative and positive UIP violations cancel each other over time (as they do in the data). In turn, the long-run dynamics of the model, and hence the unconditional macro moments, are also not too far from the model where UIP holds at all horizons.

The key is the behavior of the *level* of the real exchange rate. Note that we can write the level of the real exchange rate as:

$$q_t = \sum_{k=0}^{\infty} E_t(r^*_{t+k} - r_{t+k}) - \sum_{k=0}^{\infty} E_t(\lambda_{t+k+1})$$

Due to the reversal in UIP violations at longer horizons, at the benchmark calibration (and similarly in my empirical estimates) we have $\sum_{k=0}^{\infty} E_t(\lambda_{t+k+1}) \approx 0$, hence

$$q_t \approx \sum_{k=0}^{\infty} E_t(r^*_{t+k} - r_{t+k}) = q_t^{UIP},$$

meaning that the behavior of the real exchange rate is roughly equal to the behavior under a model where UIP holds – see also the resulting impulse response functions in Figure 4. This is why most of the moments of the model remain relatively unchanged.

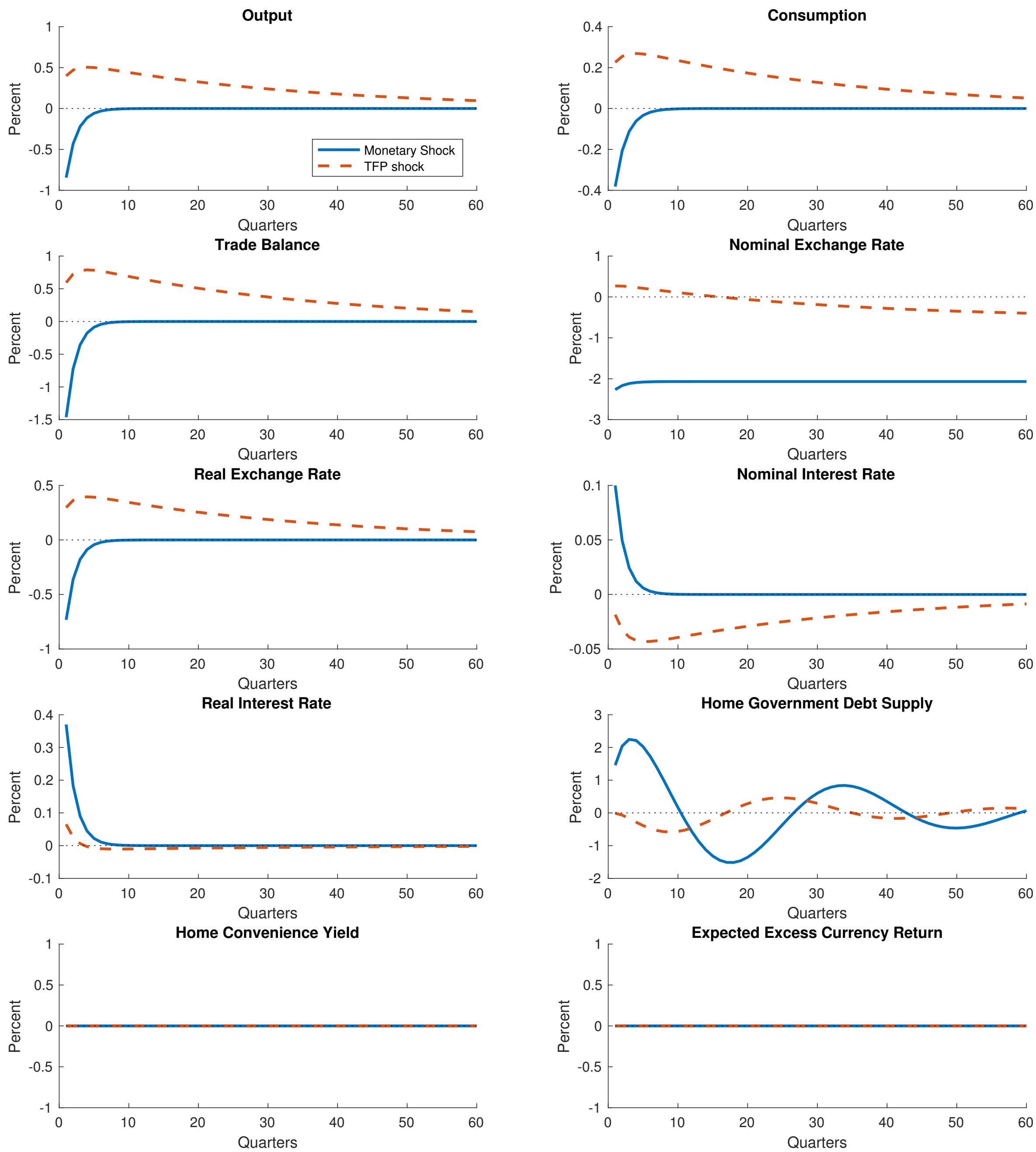
This is an important result in itself, as it implies that macroeconomic analysis can continue to rely on standard first-order models where UIP holds without much loss.

Table 1: Unconditional Moments

	Data	Benchmark Model	No Convenience Yield
Standard Deviations			
Δs_t	5.74	2.96	3.23
Δq_t	5.67	1.15	1.26
$i_t - i_t^*$	0.75	0.32	0.21
$r_t - r_t^*$	1.09	0.74	0.61
$E_t(\lambda_{t+1})$	0.48	0.31	0
Δy_t	0.78	1.06	1.18
Δc_t	0.62	0.45	0.53
$\Delta(b_t^g/y_t)$	3.15	2.49	2.65
i_t	0.84	0.44	0.30
r_t	0.91	0.76	0.66
Autocorrelations			
q_t	0.93	0.91	0.88
$i_t - i_t^*$	0.74	0.81	0.68
$r_t - r_t^*$	0.14	0.63	0.5
$E_t(\lambda_{t+1})$	0.74	0.95	0
Δy_t	0.23	-0.18	-0.20
Δc_t	0.43	-0.11	-0.14
$\Delta(b_t^g/y_t)$	0.34	0.42	0.37
i_t	0.86	0.88	0.78
r_t	0.27	0.67	0.56
Correlation with Real Exchange Rate			
$\rho(\Delta q_t, \Delta y_t)$	-0.03	0.95	0.95
$\rho(\Delta q_t, \Delta y_t - \Delta y_t^*)$	0.07	0.99	1
$\rho(\Delta q_t, \Delta c_t - \Delta c_t^*)$	-0.05	0.99	1
$\rho(\Delta q_t, \Delta s_t)$	0.98	0.84	0.87
$\rho(q_t, r_t - r_t^*)$	-0.17	-0.44	-0.51
Cross-country correlations			
$\rho(\Delta y_t, \Delta y_t^*)$	0.42	-0.82	-0.80
$\rho(\Delta c_t, \Delta c_t^*)$	0.31	0.64	0.66
$\rho(i_t, i_t^*)$	0.68	0.74	0.73
$\rho(r_t, r_t^*)$	0.49	0.52	0.57

Notes: Standard deviations are expressed in percentage terms. The data on domestic variables is for the US, the data on international variables is for the US against the other countries in the sample.

Figure 4: Impulse Response Functions



D.3 Interest Rates Across Different Types of Assets

It seems reasonable to think that some assets, like Treasuries, tend to have bigger convenience yields than other short-term assets, like say inter-bank loans. Does the model then imply that the interest rate differential (across countries) on Treasuries would behave very differently than the interest rate differentials of other, less liquid assets? That would be a potential concern, because in the data interest rate differentials across countries behave similarly, no matter what type of short-term rate one uses.

Re-assuringly, the model has no such counter-factual implications. In the model, the primary difference between different types of interests rates is in their *level*, where the interest rate of an asset with a lower convenience yield is generally higher, but the overall dynamics of interest rates across different types of assets is remarkably similar. In particular, the interest rate of a hypothetical asset that has *no* convenience yield, call it $\tilde{i}_t \equiv -E_t(\hat{M}_{t+1})$, has almost identical dynamics, and is highly correlated with the interest rate of the Treasury bill, i_t. As a result, the interest rate differentials across different types of assets are also quite similar.

For example, in the benchmark calibration the correlation between the two interest rates is 0.78, and their time series properties are quite similar – the autocorrelation of the T-bill interest rate is 0.866 and that of $\tilde{i}_t$ is 0.843. Moreover, the standard deviation of $\tilde{i}_t$ is 0.0032 and that of i_t is 0.004. And this is just a conservative lower bound on the similarity we could expect to see in the data, since there we observe assets that have lower, but still positive convenience yields (i.e. Commercial Paper). A hypothetical asset with *some* convenience yield, will look even more akin to the Treasury's in the model.

The reason for this similarity is the fundamentally negative correlation between the convenience yield and the Treasury interest rate – when the convenience yield is high, then the interest rate on the Treasury is low as investors require a lower financial compensation to hold that asset (the correlation is -0.63 in the benchmark calibration). However, this countervailing force helps make $\tilde{i}_t$ behave similarly to i_t. To see this clearly, note that the equilibrium condition linking the two interest rates in the model is

$$\tilde{i}_t = i_t + \hat{\Psi}_t^H$$

As we saw in the main text, contractionary shocks increase i_t while lowering $\hat{\Psi}_t^H$ – this is the key feature generating the UIP Puzzle, since it leads to the result that high interest rates are associated with high excess currency returns (which compensate for the low $\hat{\Psi}_t^H$). However, this exact same mechanism also leads to an increase in $\tilde{i}_t$, which generates a positive correlation between i_t and $\tilde{i}_t$. Lastly, the convenience yield is considerably less volatile than the Treasury interest rate itself – the std deviation of $\hat{\Psi}_t^H$ is only half of that of i_t. These forces together result in a high, positive correlation between i_t and $\tilde{i}_t$.

Thus, the bottom line is that the model implies that the interest rates on different types of assets, some more liquid than others, will be highly correlated and overall behave very similarly. Just like what we observe in the data.

D.4 Long-term Bonds

It is well known that the UIP holds better in the "long-run". Specifically, Chinn (2006) and others have shown that 5-year (and longer) excess currency returns display smaller UIP deviations, than the typical estimates of the UIP Puzzle in short-term bonds.

In this section, I show that the model can match this observation. I will first show that this is the case even if we make the strong (and counter-factual) assumption that long-term bonds are perfect substitutes for short-term bonds in terms of liquidity, and hence earn the *same* convenience yield. This is enough to conclude that in reality, since long-term bonds most likely have lower convenience yields than short-term bonds, there is even less reason to expect UIP deviations in long-term bonds.

The key empirical result centers on the regression

$$s_{t+N} - s_t + R_t^{*,(N)} - R_t^{(N)} = \alpha^{(N)} + \beta^{(N)}(R_t^{(N)} - R_t^{*,(N)}) + \varepsilon_{t+N}^{(N)}$$

where the $R_t^{(N)} = N * i_t^{(N)}$ is the cumulative interest rate on a $N-$period bond ($i_t^{(N)}$ is the yield on the N-period bond). The left-hand variable is the excess return on N-period foreign bond over a N-period home bond when both are held to maturity. It turns out, that while $\beta^{(N)}$ is large and significantly negative for $N \leq 1$ years, the estimates are smaller and often insignificant for $N \geq 5$ years. In other words, long-term bond returns appear to be equalized across countries, even though the short-term bonds display a clear violation of UIP.

In the model, this observation is trivially true if we assume that long-term bonds do not offer any of the convenience benefits of short-term bonds. But the point of this section is to show that the relation will still hold, *even if* long-term bonds are perfect substitutes for short-term bonds. The intuition is that multi-period excess currency returns offset the sum of expected convenience yield differentials that accrue throughout the life of the bond. So if we are looking at a 5-year bond, then the 5-year cumulative excess return will equal the expected sum of convenience yield differentials for the next 5 years. Crucially, the convenience yield differential switches signs at longer horizons (recall that this is what generates the reversal in UIP violations), and thus for long-term bonds (in particular 7+ years) the sum of expected convenience yield differentials is roughly zero. Thus, long-term excess currency returns end up being equalized, even though the short-term excess returns are not, due to the cyclical movements in the convenience yield differential analyzed in the main body of the text.

To make this concrete, assume that the convenience benefit is again derived from a similar transaction cost function $\Psi(c_t, m_t, b_t^T, b_t^{*,T})$, where b_t^T this time is the total amount of home bonds, across all maturities, in the agent's portfolio:

$$b_t^T = b_t^{(1)} + b_t^{(2)} + \dots$$

and $b_t^{*,T}$ is similarly the total amount of foreign bonds owned. Thus, the short-term bonds are no longer special relative to the longer maturity ones – they all enter equivalently in the transaction costs function.

The resulting Euler equation for 1-period bonds is the same as before:

$$E_t(\Delta s_{t+1} + i_t^* - i_t) = \hat{\Psi}_t^H - \hat{\Psi}_t^F \tag{12}$$

where $\hat{\Psi}_t^H$ and $\hat{\Psi}_t^F$ are the log-linearized home and foreign convenience yields. Note that the convenience yields on bonds across all maturities are the same, because the derivatives of the transaction cost $\Psi(.)$ in terms of different maturities are equal. That is, all bonds of the same currency denomination are equivalent to each other in terms of liquidity.

We can derive a similar Euler equation for an arbitrary N-periods to maturity bond:

$$E_t(\Delta s_{t+1} + \hat{p}_{t+1}^{*,(N-1)} - \hat{p}_t^{*,(N)} - (\hat{p}_{t+1}^{(N-1)} - \hat{p}_t^{(N)}) = \hat{\Psi}_t - \hat{\Psi}_t^*$$

where $\hat{p}_t^{(N)}$ is the (log-linearized) price of the N period (zero-coupon) bond. The cumulative interest rate payments of the bond are $R_t^{(N)} = N * i_t^{(N)} = \frac{1}{p_t^N}$, and hence

$$E_t(\Delta s_{t+1} + \hat{R}_{t+1}^{*,(N)} - \hat{R}_{t+1}^{(N)} - (\hat{R}_t^{*,(N-1)} - \hat{R}_t^{(N-1)})) = \hat{\Psi}_t - \hat{\Psi}_t^*$$

Solving recursively for $\hat{R}_t^{*,(N-1)} - \hat{R}_t^{(N-1)}$, and substituting it back and leads to

$$E_t(\Delta s_{t+1} + \hat{R}_{t+1}^{*,(N)} - \hat{R}_{t+1}^{(N)}) = E_t \sum_{k=0}^{N-1} (\hat{\Psi}_{t+k} - \hat{\Psi}_{t+k}^*)$$

This is very intuitive – the excess return on a carry trade (held to maturity) is the sum of expected future convenience yield differentials. Thus, when operating with 1-period bonds we have equation (12), so that only the current convenience yield matters, but when we consider long-term bonds then it is the whole path of expected convenience yield differentials. And since in the model the convenience yield differential changes signs at longer horizons (see Figure 4 in the main text for example), the sum $E_t \sum_{k=0}^{N-1} (\hat{\Psi}_{t+k} - \hat{\Psi}_{t+k}^*)$ in fact grows smaller for higher N. Due to their cyclical dynamics that underpin the key results of the model, the convenience yields further into the future cancel out the shorter-horizon ones. In particular, in the benchmark calibration of the model, the sum at horizons of 7 years or more is roughly zero, which matches the data well.

D.5 Term-Structure Effects in UIP violations in the data

We can further examine the empirical evidence on the UIP violations, and decomposes the documented UIP violations into a pure exchange rate effect and a term-structure effect due to violations in the expectations hypothesis (EH) of the interest rate term-structure. The results show that the pure exchange rate component is the primary driver of the estimated UIP violations and their changing nature. This is another reason for why abstracting away from long-term bonds and term structure effects, as I do in the model, is unlikely to be important.

According to the EH, cumulative long-term interest rates are equal to the sum of expected future short-rates over the duration of the long-term interest rate. This implies that a zero coupon n-month bond's cumulative interest rate, $R_t^{(n)}$, is given by

$$R_t^{(n)} = \sum_{k=0}^{n-1} E_t(i_{t+k}),$$

where, as before, i_t is the 1-month interest rate at time t. We can then use this relation to back out risk-neutral expectations of future short-rates from the term-structure itself. Let $i_{t,t+k}$ be the time-t risk-neutral expectation of the 1-month interest rate at time $t+k$, also known as the *forward* interest rate at time t, and note that this is given by the difference in interest rates of a $(k+1)$-months bond and a k-months bond:

$$i_{t,t+k} = R_t^{(k+1)} - R_t^{(k)},$$

Clearly, $i_{t,t} = i_t$, but as has been shown extensively in the bond literature, the EH hypothesis fails at longer horizons (e.g. Campbell and Shiller (1991)), and the forecast errors

$$\eta_{t,t+k} = i_{t,t+k} - i_{t+k}$$

are forecastable by today's (time t) short-rate.

To see how this could affect currency return forecasts, add and subtract the forward interest differential $i^*_{t,t+k} - i_{t,t+k}$ from the excess currency return λ_{t+k+1} to obtain

$$E_t(\lambda_{t+k}) = E_t(\Delta s_{t+k} + i^*_{t,t+k-1} - i_{t,t+k-1}) + E_t(\underbrace{i^*_{t+k-1} - i_{t+k-1} - (i^*_{t,t+k-1} - i_{t,t+k-1})}_{\eta^*_{t,t+k-1} - \eta_{t,t+k-1}}). \tag{13}$$

Forecastability in excess currency returns could arise from either of the two components above. The first piece measures how well exchange rates offset forward interest rates, and captures the pure exchange rate effect. In essence, it is the expected excess currency return in a world where the EH holds.[2] The second component measures the forecastability of interest-rate excess returns themselves, which captures the term-structure anomaly effect. Next, I decompose the forecastability of excess currency returns into these two components.

To do so, I construct a zero-coupon term-structure of interest rate differentials by using the forward discount at maturities of up to a year, and data on interest rate swaps from Bloomberg for longer maturities. Data on long-maturity interest rates is only available starting in 1990, and the shorter time-series leads me to drop the Euro-legacy currencies from the benchmark results, because they are left with less than 10 years of data. This leaves me with a data on 10 currencies for the period 1990-2013, for which I compute the two components in (13) and run separate forecasting regressions on each

$$s_{t+k} - s_{t+k-1} + i^*_{t,t+k-1} - i_{t,t+k-1} = \alpha_{j,k} + \delta_k(i_{j,t} - i^*_{j,t}) + \nu_{j,t+k}$$

$$i^*_{t+k} - i_{t+k} - (i^*_{t,t+k} - i_{t,t+k}) = a_{j,k} + \theta_k(i_{j,t} - i^*_{j,t}) + v_{j,t+k}$$

to estimate δ_k and θ_k, which by construction sum up to the original UIP coefficients β_k

$$\beta_k = \delta_k + \theta_k.$$

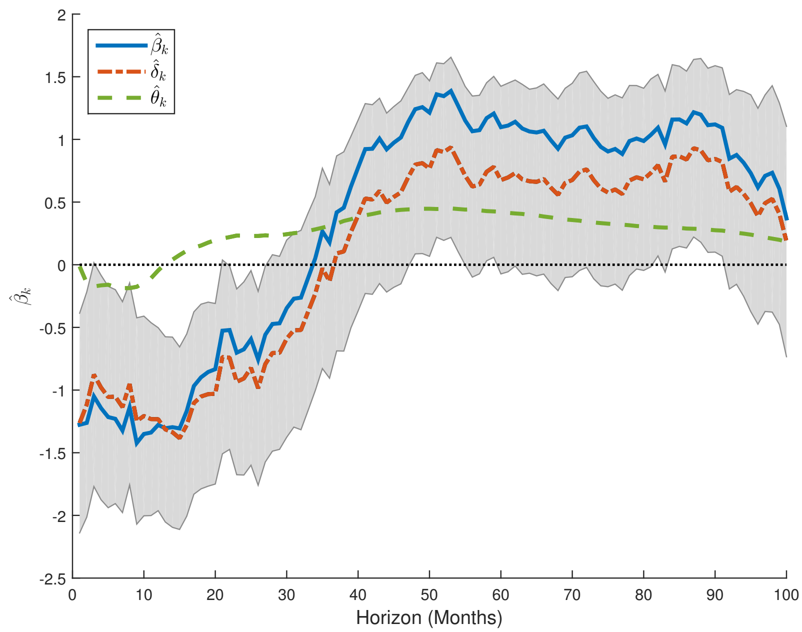
Thus, these two series of estimates decompose the UIP violations into a pure exchange rate effect, δ_k, and a term-structure effect, θ_k. The results are plotted in Figure 5, where the blue line represents the original $\hat{\beta}_k$ estimates (but now estimated on the smaller data set for comparison purposes), the red dash-dot line plots $\hat{\delta}_k$ and the green dashed line plots $\hat{\theta}_k$. The shaded region represents the 95% confidence interval around the estimates of δ_k.

The results show that the exchange rate behavior is the primary driver of the cyclicality in excess currency returns. The $\hat{\delta}_k$ estimates are statistically significant, track $\hat{\beta}_k$ closely and display a very similar pattern across horizons, where they start out negative, and then turn positive at the same time as $\hat{\beta}_k$. In terms of overall magnitudes, the $\hat{\delta}_k$ coefficients account for virtually all of the negative UIP violations at horizons of less than 36 months, and for more

[2]This is not a purely theoretical construct, this return can be obtained by going long the excess return on a foreign $k+1$ months bond and short the excess return on a k-months foreign bond:

$$\Delta s_{t+k+1} + i^*_{t,t+k} - i_{t,t+k} = s_{t+k+1} - s_t + R_t^{(k+1)*} - R_t^{(k+1)} - (s_{t+k} - s_t + R_t^{(k)*} - R_t^{(k)})$$

Figure 5: UIP Violations Decomposition



than two-thirds of the positive UIP violations at longer horizons.[3] On the other hand, while the term-structure effects are also non-zero and switch from negative to positive, their timing is quite different and the magnitude is much smaller. Thus, the results point to exchange rate behavior as the most important driver of the changing nature of UIP violations, with term-structure effects playing only a secondary role. As such, modeling short-term bonds only is sufficient to understand the first-order features of the puzzle.

D.6 Empirical Debt Dynamics

Cyclical debt dynamics are an integral part of the mechanism, and in this section I verify that the data displays non-monotonic dynamics similar to the model. I focus on US government debt, because it is available at a quarterly frequency for the whole sample period, while foreign government debt series are available only at the annual level before 1991.

I estimate the impulse response of government debt using the same Jorda projection methods as the ones used to estimate the dynamics of the excess currency return. So I run a series of regressions indexed by k

$$b_{t+k} = \mu + \beta_k b_t + \varepsilon_{t+k},$$

where b_t is the log of US federal debt held by the public (variable FYGFGDQ188S in FRED),

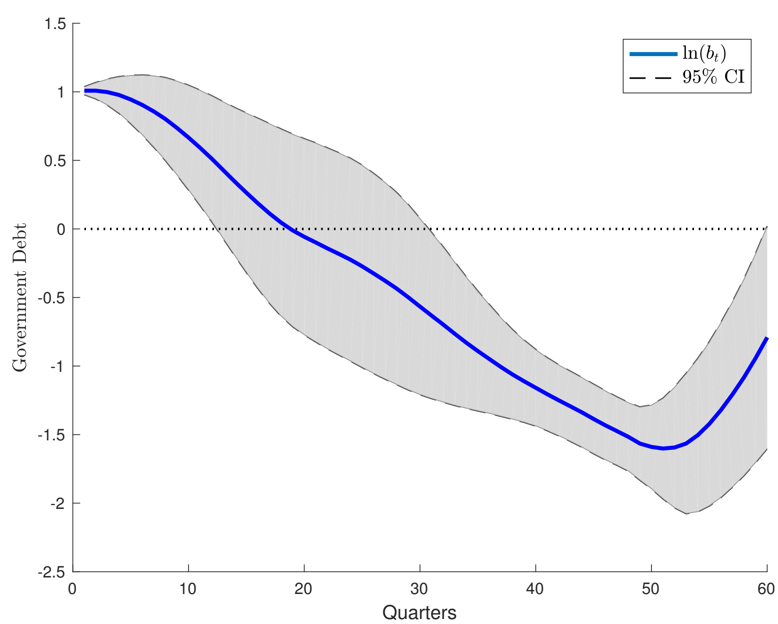
[3]While the $\hat{\delta}_k$ estimates barely miss the 95% significance cut-off at 60-80 month horizons, they are significant at the 90% level at all horizons.

after removing an exponential time trend.[4] Lastly, the data is quarterly, hence the index k controls the number of quarters ahead each forecast is made for. As before, the sequence of β_k forms an estimate of the impulse response of government debt to an increase in today's debt level.

I estimate the dynamics of debt in this way for two reasons. First, I want to remain agnostic about the source of shocks, and rather than try to identify specific structural shocks, I want to estimate the overall dynamics government debt. As we saw in section **??**, the source of shocks does not matter in the model – due to the interaction of monetary and fiscal policy, the dynamics of government debt are determined by complex roots, and thus display cyclicality regardless of the shock. Second, the key motivating empirical fact of cyclical excess currency returns are also estimated via the same Jorda projections method.

The resulting IRF is plotted in Figure 6 below. As we can see, in the data US debt dynamics display the type of cyclicality implied by the model. Debt is highly persistent and an increase in debt lasts for several years. Importantly, on the way down debt does not converge monotonically, but dips significantly below its long-run mean before converging. In other words, it displays the type of cyclicality implied by the model and also observed in the excess currency return in the data. Moreover, the timing of crossing zero is similar to the one observed in currency returns – debt falls below steady state after about 4-5 years, which is roughly the same as with the currency returns.

Figure 6: Debt Impulse Response



[4]Moreover, the results remain qualitatively the same when using VARs and structural identification schemes. Results are also unchanged if we use Debt-to-GDP ratio instead of detrended debt in levels.

D.7 Steady State Implications

At the zero-inflation steady state, the Euler equations for domestic and foreign bonds imply that the interest rate differential and the steady state excess currency returns are given by

$$i - i^* = \frac{1}{\beta}(\Psi^F - \Psi^H)$$

$$(1+i^*)\frac{S'}{S} - (1+i) = \underbrace{\frac{1}{\beta}(\Psi^H - \Psi^F)}_{\text{Convenience Yield Differential}}$$

Thus, if there are cross-sectional differences in the steady state convenience values of assets denominated in different currencies, this will drive a corresponding difference in their steady state interest rates as well. Importantly, we would expect that a higher convenience yield differential is associated with a lower interest rate differential. In addition, differences in the convenience yields will also lead to a non-zero steady state excess currency return. When the home convenience yield is higher than the foreign one, the foreign currency will be compensated through a positive excess return, in order to keep investors indifferent between home and foreign bonds.

Hence, the model can explain the Hassan and Mano (2018) evidence that a big portion of carry trade returns are due to persistent cross-sectional differences in currencies and unconditional premia, and not time-variation in conditional premia. For example, the model would imply that part of the reason why the Japanese yen is consistently a funding currency and the Australian dollar is consistently an investment currency, is because the Japanese yen is a major international reserve currency while the Australian dollar is not. As such, the yen earns a higher convenience yield on average, and thus has a relatively lower interest rate and negative excess returns versus the Australian dollar.

Thinking about the drivers of the unconditional premia of carry trades is an interesting question, but is distinct from the primary motivation of this paper – the cyclical nature of UIP violations. To understand the UIP regression evidence, and its changing nature at different horizons, one needs to understand the equilibrium dynamics of the conditional excess currency returns. To this end, in this paper I focus on the symmetric steady state where $\Psi^H = \Psi^F$ in order to isolate the effect of the time-variation in the convenience yield. Analyzing the behavior the model around asymmetric steady states is an interesting avenue for future work. For work in this direction, please see Chahrour and Valchev (2017) who provide a model with multiple steady states, including asymmetric ones.

E Debt and Excess Currency Returns Extra Results

Including post-2007 data: Table 2 below re-estimates the regression specifications of Section 6.1 in the main text,

$$\lambda_{j,t+1} = \alpha_j + \beta(i_t - i^*_{j,t}) + \gamma \ln(\frac{Debt^{US}_t}{GDP^{US}_t}) + \gamma^* \ln(\frac{Debt^*_{jt}}{GDP^*_{jt}}) + \delta \ln(\frac{CP^{US}_t}{GDP^{US}_t}) + Controls + \varepsilon_{j,t+1},$$

by also including the post-2007 period. The results remain similar to the benchmark specification in the main text – namely the coefficient on US government debt is large, negative and significant, and while smaller in magnitude, the coefficient on foreign government debt also generally shows up as statistically significant.

Table 2: Excess Currency Returns and Debt: 1991-2013

	(1)	(2)	(3)	(4)	(5)
$(i_t - i^*_t) * 100$	-1.41***	-1.52***	-0.77	-1.38	-1.23
	(0.46)	(0.46)	(0.49)	(0.94)	(0.96)
$\ln(\frac{Debt}{GDP})$		-0.47	-3.02***	-3.85***	-3.67***
		(0.39)	(1.06)	(1.32)	(1.36)
$\ln(\frac{Debt^*}{GDP^*})$		0.17	0.25**	0.13	0.21*
		(0.11)	(0.12)	(0.13)	(0.11)
$\ln(\frac{CP}{GDP})$			-2.47***	-3.20***	3.05***
			(0.94)	(1.28)	(1.30)
$\frac{NFA}{GDP}$					1.29
					(1.12)
$\widehat{VIX}$					0.25
					(0.40)
.					
KVJ2012 Controls	No	No	No	Yes	Yes
# Currencies	18	18	18	15	15
Fixed Effects	Yes	Yes	Yes	Yes	Yes

Estimates with Driscoll and Kraay (1998) standard errors robust to heteroskedasticity, serial correlation and cross-equation correlation. The debt stock variables are exponentially detrended. ***, ** and * denote significance at the 1%, 5% and 10% level respectively. The LHS variable, excess currency returns, is expressed in terms of percent, i.e. the LHS is $\lambda_{j,t+1} * 100$

Quarterly Frequency Results: Table 3 below re-estimates the regression specifications of Section 6.1 in the main text,

$$\lambda^{3m}_{j,t+1} = \alpha_j + \beta(i^{3m}_t - i^{3m,*}_{j,t}) + \gamma \ln(\frac{Debt^{US}_t}{GDP^{US}_t}) + \gamma^* \ln(\frac{Debt^*_{jt}}{GDP^*_{jt}}) + \delta \ln(\frac{CP^{US}_t}{GDP^{US}_t}) + Controls + \varepsilon_{j,t+1},$$

by using quarterly frequency data only. To match the data frequency, the excess currency returns and the interest rate differentials are for 3-month horizons. The overall results and significance is very similar to the main specifications reported in the main text. Crucially, the coefficient on US government debt to GDP remains large, negative and significant, in line with the benchmark results. The magnitude of the coefficient estimate here is about 3 times as large as the benchmark estimates, as should be expected given that the left-hand side here is 3-month excess returns, whereas it is 1-month excess returns in the daily frequency regressions.

Table 3: Excess Currency Returns and Debt, Quarterly Frequency: 1991 -2007

	(1)	(2)	(3)	(4)	(5)
$(i^{3m}_t - i^{3m,*}_t) * 100$	-1.48**	-1.37	0.17	0.76	1.73
	(0.72)	(0.95)	(1.96)	(2.29)	(2.23)
$\ln(\frac{Debt}{GDP})$		-1.43	-18.06**	-16.34*	-23.99**
		(2.35)	(8.55)	(9.36)	(11.74)
$\ln(\frac{Debt^*}{GDP^*})$		0.39	-0.23	0.02	0.53
		(0.69)	(0.54)	(0.51)	(0.51)
$\ln(\frac{CP}{GDP})$			-11.20*	-11.01	-12.23
			(6.14)	(7.06)	(8.72)
$\frac{NFA}{GDP}$					21.58
					(17.04)
VIX					5.34**
					(2.32)
.					
KVJ2012 Controls	No	No	No	Yes	Yes
# Currencies	10	10	10	10	10
Fixed Effects	Yes	Yes	Yes	Yes	Yes

Estimates with Driscoll and Kraay (1998) standard errors robust to heteroskedasticity, serial correlation and cross-equation correlation. The debt stock variables are exponentially detrended. ***, ** and * denote significance at the 1%, 5% and 10% level respectively.The LHS variable, excess currency returns, is expressed in terms of percent, i.e. the LHS is $\lambda_{j,t+1} * 100$

Utilizing longer US data series: Table 4 below re-estimates the regression specifications of Section 6.1 in the main text,

$$\lambda_{j,t+1} = \alpha_j + \beta(i_t - i^*_{j,t}) + \gamma \ln(\frac{Debt^{US}_t}{GDP^{US}_t}) + \gamma^* \ln(\frac{Debt^*_{jt}}{GDP^*_{jt}}) + \delta \ln(\frac{CP^{US}_t}{GDP^{US}_t}) + Controls + \varepsilon_{j,t+1},$$

by making use of the longer availability of US data for government debt and commercial paper. Thus, the data for those regressions starts in 1984, the earliest availability of USD commercial paper data. By necessity, the regressions exclude foreign debt and the VIX index due to the lack of data going back to 1984 – the VIX starts in 1990 and the foreign debt data in 1991, so for those specifications refer to the main text results in Section 6.1. The controls vector, still includes foreign stock market volatility and yield slope, but once those are included the sample shrinks to 15 countries, out of 18 total.

All results remain the same as before, both quantitatively and qualitatively – there is a large and significant negative coefficient values on US debt, even in this longer sample.

Table 4: Excess Currency Returns and Debt: 1984-2007 (US debt only)

	(1)	(2)	(3)	(4)	(5)
$(i_t - i^*_t) * 100$	-1.07***	-1.13***	-1.02**	-1.13**	-1.17**
	(0.41)	(0.41)	(0.45)	(0.47)	(0.47)
$\ln(\frac{Debt}{GDP})$		-1.49***	-1.75***	-1.83***	-2.17***
		(0.58)	(0.65)	(0.69)	(0.77)
$\ln(\frac{CP}{GDP})$			-0.35	0.03	0.99
			(0.43)	(0.48)	(1.03)
$\frac{NFA}{GDP}$					3.38
					(3.23)
.					
KVJ2012 Controls	No	No	No	Yes	Yes
# Currencies	18	18	18	15	15
Fixed Effects	Yes	Yes	Yes	Yes	Yes

Estimates with Driscoll and Kraay (1998) standard errors robust to heteroskedasticity, serial correlation and cross-equation correlation. The debt stock variables are exponentially detrended. ***, ** and * denote significance at the 1%, 5% and 10% level respectively. The LHS variable, excess currency returns, is expressed in terms of percent, i.e. the LHS is $\lambda_{j,t+1} * 100$

F UIP Violation Reversals and Monetary and Fiscal Policy

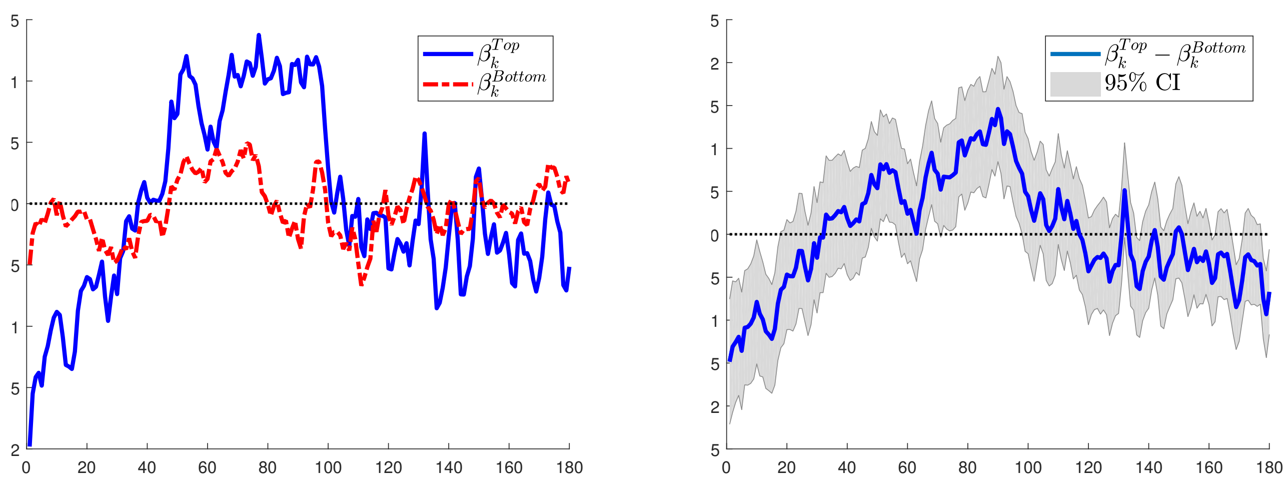
Another important feature of the model is the key role played by the interaction of monetary and fiscal policy. The model predicts that we should see clear reversals in the UIP violations only for countries that have both (i) active monetary policy and (ii) sluggish fiscal policy, and in this section I verify this in the data. This analysis is related to Bansal and Dahlquist (2000) who find that countries with higher and more volatile inflation display significantly lower violations of the classic, short-horizon UIP condition, and reason that this evidence calls for a mechanism that has an explicit role for monetary and fiscal policy. I extend their work by showing that there is also a strong cross-sectional link between monetary and fiscal policy and the reversal of UIP violations at longer horizons, as predicted by the model.

I examine this relationship in the data by first sorting currencies on their monetary policy stance, and then further sorting on their tax policy sluggishness. For completeness, I consider four different proxies for the monetary stance of a country. In addition to the two proxies used in Bansal and Dahlquist (2000), average inflation and the standard deviation of inflation, I use the Central Bank Independence Index (CBI) of Grilli et al. (1991) (updated with recent data by Arnone et al. (2007)), and the degree of capital controls, as measured by the Chinn and Ito (2006) index.[5] Since the proxies are generally only available at a low frequency, I focus on exploiting the cross-sectional dimension of the data. For each currency, I compute the corresponding average value for each proxy (e.g. average CBI for the UK over 1976-2013 and etc.), and then for each proxy I sort the currencies into two bins – high and low. Finally, I find the intersection of all the top bins, which yields five countries (Canada (CAD), Germany (DEM), the Netherlands (NLG), Switzerland (CHF) and the UK (GBP)) that score in the top half in all measures of monetary policy independence. And similarly obtain the intersection of the bottom bins, which yields (Ireland (IEP), Italy (ITL), Spain (ESP), Portugal (PTL)). Then I re-estimate the series of UIP regressions at different horizons, eq. (2) in the main text, for both sets separately and compare the results.

Figure 7 plots the estimates and shows a remarkably consistent message. In panel a) we see that currencies with high monetary independence display a much more pronounced evidence of cyclicality in UIP violations, and generally exhibit a larger magnitude of UIP violations at all horizons. Panel b) shows that the difference between the two estimates, $\beta_k^{Top} - \beta_k^{Bottom}$, is in fact statistically significant (at the 5% level). Thus, currencies with a more independent monetary policy do not only display larger UIP violations at short-horizons,

[5]Capital controls are commonly used as a *de facto* measure of CB independence – see for example Alesina and Tabellini (1989), Drazen (1989), Grilli and Milesi-Ferretti (1995), and Bai and Wei (2000)

Figure 7: UIP Violations and Monetary Policy



(a) Top vs Bottom Third

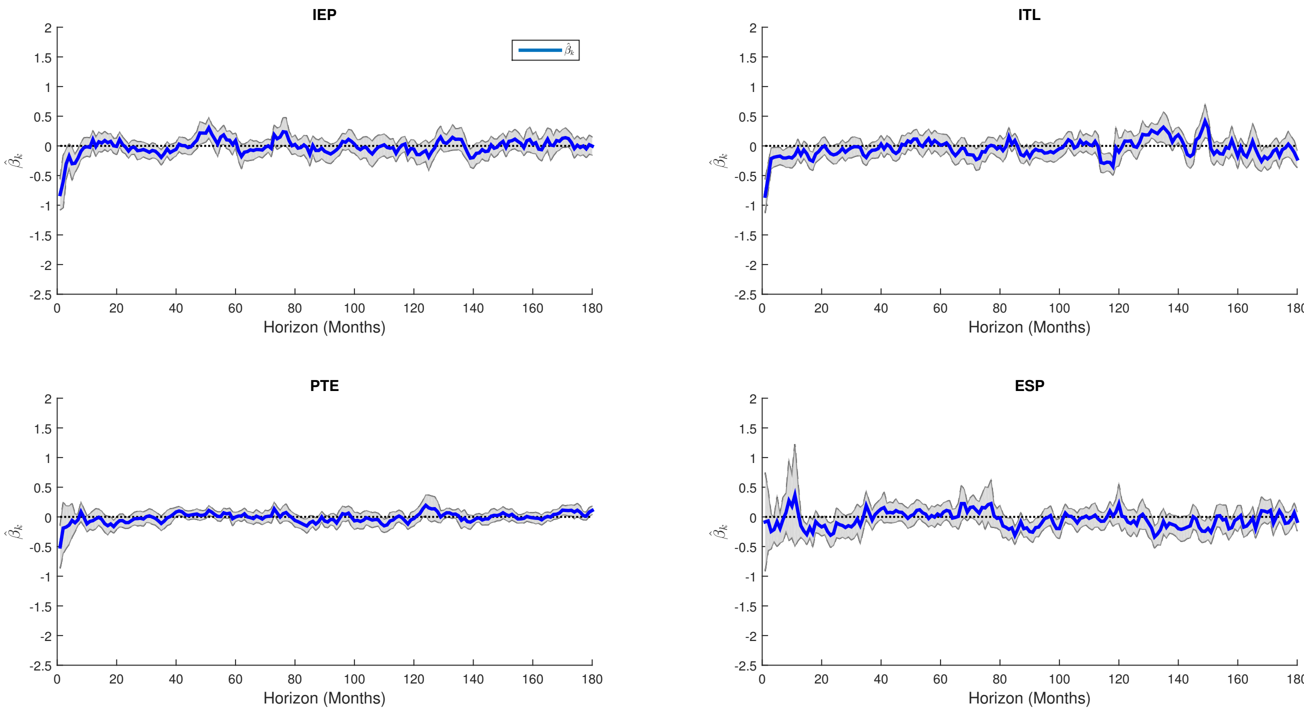
(b) $\beta_k^{Top} - \beta_k^{Bottom}$

but also stronger evidence of a reversal in their direction at longer horizons.

However, since the US scores high in all four proxies, one leg of each currency pair displays strongly independent monetary policy throughout the whole sample (recall that all currencies are quoted against the dollar). Since according to the model this is a necessary condition for UIP reversals to occur, it is interesting to also consider results where the base currency has low monetary independence. To do so, I use the set of currencies that are in the bottom bin according to all proxies (IEP, ITL, ESP, PTE) as alternative base currencies, and construct four different sets of currency pairs (e.g. ITL-AUD, ITL-ATS, ...). This gives me four data sets of 18 currencies each, that I then use to re-estimate the initial set of regressions in eq. (2) in the main text. The results are plotted in Figure 8 and are quite striking – in all four plots the UIP violations exhibit virtually no evidence of a reversal. Thus, it appears that the cyclicality in UIP violations is indeed associated with strong and independent monetary policy.

The above results are evidence that a hawkish monetary policy is a necessary condition for reversals in UIP violations, but what about fiscal policy? To answer this question, I now focus on the subset of currencies that have hawkish monetary policy (CAD, DEM, NLG, CHF, GBP and USD) and further sort them on their fiscal policy in two ways. First, I compute the autocorrelation of the growth in public debt (both in levels and relative to GDP), which will be positive when taxes are relatively sluggish and debt displays non-monotonic dynamics (as evidenced by the moments in the quantitative model). Only three countries have positive such autocorrelations – CAD, GBP and USD. Second, I estimate the tax policy rule posited by the model, compute the implied threshold value $\underline{\rho}(\kappa_b)$ as per Lemma 2 and check which

Figure 8: UIP Regressions, 1 to 180 months



countries have ρ_τ estimates above that threshold (and thus would be predicted to display cyclical dynamics). By this second criterion, we would again expect to see UIP violations reversals for CAD, GBP and USD (and to a lesser extent DEM).

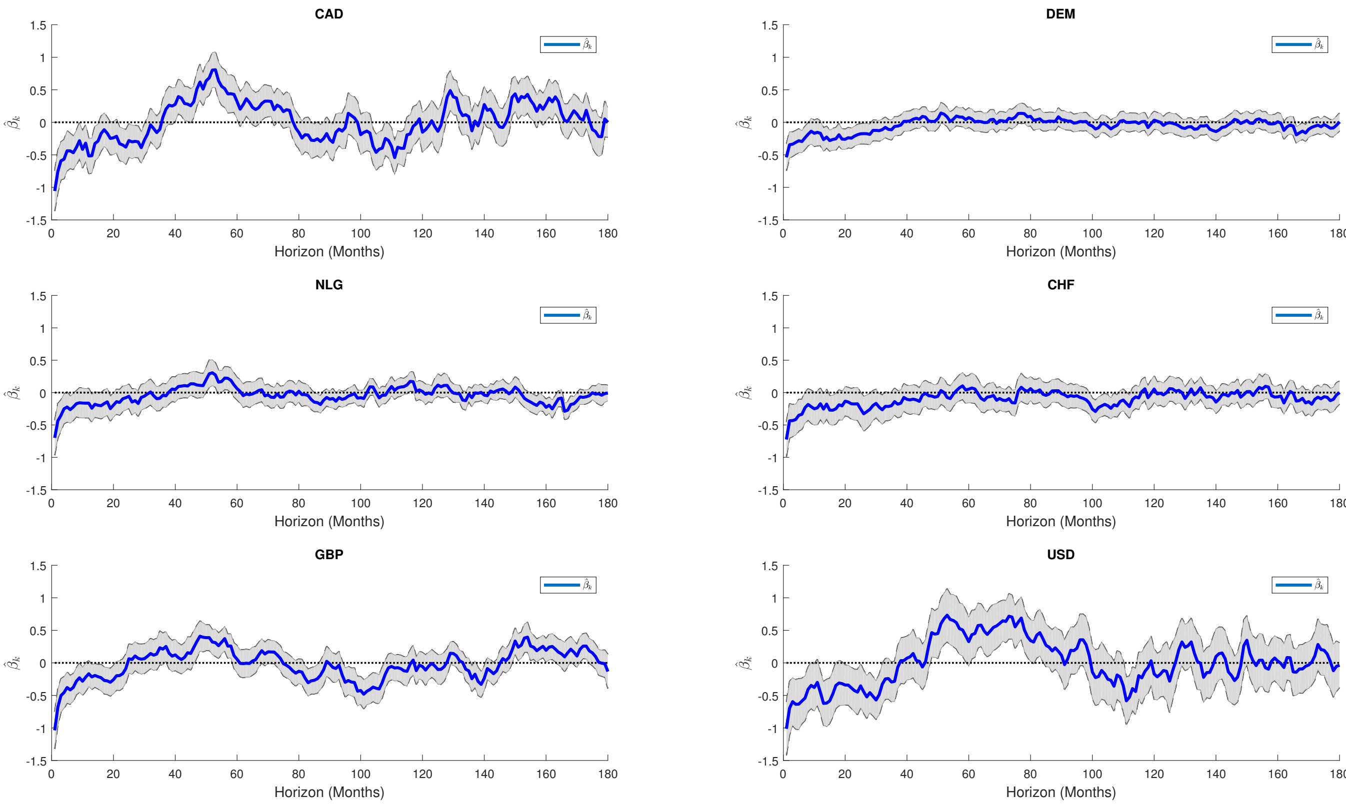
To check these predictions, I now compute a version of Figure 8 where I use the six currencies with strong monetary policy as alternative base currencies. I plot the results in Figure 9, which shows that the predictions of the model are borne out by the data. It is not the case that *all* of the six currencies display cyclicality in the UIP violations. Only the currencies with sluggish tax policies (CAD, GBP and USD) clearly do so, which supports the model's implication that monetary policy is only a necessary, but not sufficient condition. Crucially, it is the interaction of both an active monetary policy and a sluggish fiscal policy that is associated with cyclical movements in UIP violations, just as predicted by the model.

References

Alesina, Alberto and Guido Tabellini, "External debt, capital flight and political risk," *Journal of International Economics*, 1989, *27* (3), 199–220.

Arnone, Marco, Bernard Laurens, Jean-François Segalotto, and Martin Sommer, "Central bank autonomy: lessons from global trends," *IMF Working Papers*, 2007, pp. 1–53.

Figure 9: UIP Regressions, 1 to 180 months



Bai, Chong-En and Shang-Jin Wei, "Quality of bureaucracy and open-economy macro policies," Technical Report, National bureau of economic research 2000.

Bansal, Ravi and Magnus Dahlquist, "The forward premium puzzle: different tales from developed and emerging economies," *Journal of international Economics*, 2000, *51* (1), 115–144.

Campbell, John Y and Robert J Shiller, "Yield spreads and interest rate movements: A bird's eye view," *The Review of Economic Studies*, 1991, *58* (3), 495–514.

Chahrour, Ryan and Rosen Valchev, "International Medium of Exchange – Privilege and Duty," Technical Report, Boston College Working Paper Series 2017.

Chinn, Menzie D, "The (partial) rehabilitation of interest rate parity in the floating rate era: Longer horizons, alternative expectations, and emerging markets," *Journal of International Money and Finance*, 2006, *25* (1), 7–21.

— **and Hiro Ito**, "What matters for financial development? Capital controls, institutions, and interactions," *Journal of Development Economics*, 2006, *81* (1), 163–192.

Drazen, Allan, "Monetary policy, capital controls and seigniorage in an open economy," *A European central bank*, 1989, pp. 13–32.

Driscoll, John C and Aart C Kraay, "Consistent covariance matrix estimation with spatially dependent panel data," *Review of economics and statistics*, 1998, *80* (4), 549–560.

Du, Wenxin, Alexander Tepper, and Adrien Verdelhan, "Deviations from covered interest rate parity," *The Journal of Finance*, 2018, *73* (3), 915–957.

Grilli, Vittorio and Gian Maria Milesi-Ferretti, "Economic Effects and Structural Determinants of Capital Controls," *IMF Staff Papers*, 1995, *42* (3), 517–551.

_, Donato Masciandaro, Guido Tabellini, Edmond Malinvaud, and Marco Pagano, "Political and monetary institutions and public financial policies in the industrial countries," *Economic policy*, 1991, pp. 342–392.

Hassan, Tarek A and Rui C Mano, "Forward and spot exchange rates in a multi-currency world," *The Quarterly Journal of Economics*, 2018, *134* (1), 397–450.