Exchange Rates, Interest Rates, and Gradual Portfolio Adjustment\textsuperscript{1}

Philippe Bacchetta  
University of Lausanne  
Swiss Finance Institute  
CEPR

Eric van Wincoop  
University of Virginia  
NBER

March 2, 2018

\textsuperscript{1}We gratefully acknowledge financial support from the Swiss National Science Foundation and the Bankard Fund for Political Economy.
Abstract

to be written...
1 Introduction

Richard Thaler won the 2017 Nobel Prize in Economics for incorporating “psychologically realistic assumptions into analyses of economic decision-making,” according to the press release by the Nobel committee. One area in which Thaler has noticed behavior inconsistent with what he refers to as “rational efficient markets,” is the foreign exchange market. Froot and Thaler (2000) argue that “a rational efficient markets paradigm provides no satisfactory explanation for the observed results,” referring to the forward discount puzzle that high interest rate currencies tend to appreciate. Froot and Thaler (2000) refer to this puzzle as an anomaly and propose a solution. Their hypothesis is that “...at least some investors are slow in responding to changes in the interest differential.” They argue that “It may be that these investors need some time to think about trades before executing them, or that they simply cannot respond quickly to recent information.” In Bacchetta and van Wincoop (2010) we took this proposal seriously and showed that it can indeed account for the forward discount puzzle. In this paper our objective is to show that the proposal by Froot and Thaler (2000) can in fact account for a much broader set of puzzles about the relationship between interest rates and exchange rates that have been identified in the literature.

The six puzzles that we will address are:

1. delayed overshooting puzzle: a monetary contraction that raises the interest rate, leads to a period of appreciation, followed by gradual depreciation

2. forward discount puzzle (or Fama puzzle): high interest rate currencies have higher expected returns over the near future

3. predictability reversal puzzle: high interest rate currencies after some period of time will have lower expected returns

4. Engel puzzle: high interest rate currencies are stronger than implied by uncovered interest parity

5. forward guidance exchange rate puzzle: the exchange rate is more strongly affected by expected interest rates in the near future than the distant future

6. LSV puzzle: long-term bond return differentials across countries are not predictable by current interest differentials
We will argue that gradual portfolio adjustment, where agents adjust their portfolios slowly to changes in expected returns, can account for all of these puzzles.

The delayed overshooting puzzles was first documented by Eichenbaum and Evans (1995) for the US and Grilli and Roubini (1996) for other countries. It should be pointed out that that some of the subsequent studies have shown that the puzzle does not apply to all countries and time periods and the evidence depends on identification strategies. The forward discount puzzle is the most well-known puzzle on this list and continues to be a well established empirical fact. The predictability reversal puzzle, documented by Bacchetta and van Wincoop (2010) is related to the forward discount puzzle. They show that while the excess return over the next quarter is positive for higher interest rate currencies (forward discount puzzle), after about 8 quarters the quarterly excess return is negative for currencies whose current interest rate is relatively high. In other words, there is a reversal in the sign of expected excess returns. Engel (2016) confirms that this is a robust puzzle.

The Engel puzzle is developed in Engel (2016). The exchange rate can be written as the sum of all expected future interest differentials (the UIP exchange rate that applies under uncovered interest rate parity) plus the sum of all future expected excess returns. The Engel puzzle says that the sum of all expected future excess returns is negative for high interest rate currencies. In other words, the predictability reversal will ultimately dominate and investors demand a lower sum of future excess returns on currencies whose interest rate is currently high. Such currencies are therefore strong relative to what they would be under UIP. Engel (2016) argues that this evidence is inconsistent with both standard models where the excess returns are risk premia and a model with gradual portfolio adjustment. We will see that the latter is not correct and is due to a misunderstanding of models with gradual portfolio adjustment.

The forward guidance exchange rate puzzle is developed by Gali (2017). As stated above, under UIP the exchange rate is equal to the unweighted sum of all future expected interest rate differentials. This implies that changes in expected interest rates in the near future have the same effect on the exchange rate today as changes in the expected interest differential in the more distant future. However, in the data Gali (2017) finds that expectations of interest differentials in the distant future have a much smaller effect on the current exchange rate than expectations of interest differentials in the near future.
The LSV puzzle stands for the puzzle developed by Lustig, Stathopoulos and Verdelhan (2017). It says that the forward discount puzzle has no analogy in long-term bonds. While the international excess return on short-term bonds tends to be positive for currencies with a relatively high interest rate (forward discount puzzle), LSV find that this is not the case for long-term bonds. For example, the excess return of Australian long-term bonds over US long term bonds can be written as the sum of the international excess return of short term bonds plus the difference between Australian and US local excess returns of long-term over short term bonds. Lustig et al. (2017) show that while high interest rate currencies have a positive expected excess return for short-term bonds, the local excess return of long-term bonds over short term bonds tends to be lower for high interest rate currencies. They find that no-arbitrage models in international finance cannot account for this.

By now there is significant evidence of gradual portfolio adjustment in investors’ portfolios. Ameriks and Zeldes (2004) document that investors make changes to their TIAA-CREF allocations very infrequently. Using data from the Panel Study of Income Dynamics (PSID) and the Survey of Consumer Finances (SCF), Bilias et al. (2010) find widespread inertia of portfolios in response to stock market fluctuations. Brunnermeier and Nagel (2008) use PSID data to conclude that “...one of the major drivers of household portfolio allocation seems to be inertia: households rebalance only very slowly following inflows and outflows or capital gains and losses.” Mitchell et al. (2006) find that 401(k) plan participants are characterized by “profound inertia”. Duffie (2010) reviews a broad range of evidence motivating models of gradual portfolio adjustment.

There are three ways to model gradual portfolio adjustment. The most popular way in the literature is to assume that agents change their portfolio at a fixed interval, every \( T \) periods.\(^1\) This may or may not be motivated by a cost of adjusting portfolio allocation. Recently, Bacchetta and van Wincoop (2017) model it by assuming that agents change their portfolio each period with a given probability. In the context of price setting by firms, the first assumption is analogous to Taylor price setting, while the second assumption is analogous to Calvo price setting. Bacchetta and van Wincoop (2017) argue that the assumption of a constant probability of changing the portfolio has the advantage of a more smooth response to

\(^1\)For examples, see Bacchetta and van Wincoop (2010), Bogousslavky (2016), Duffie (2010), henderschott et al. (2013) and Greenwood et al. (2015).
shocks. When agents change their portfolio every $T$ periods, investors anticipate that agents who changed their portfolio at the time of a shock will change their portfolio again $T$ periods later. This anticipation significantly affects the impulse response to shocks, making it less smooth.

In this paper we will adopt a third way of modeling gradual portfolio adjustment. It takes a bit of a shortcut relative to the alternatives discussed above, but this comes at a significant gain of analytical tractability that delivers key insights. We will assume that agents can change their portfolio allocation each period, but face a quadratic cost of changing their portfolio.\footnote{\cite{VayanosWoolley2012} also introduce a cost of changing portfolios to model portfolio inertia. They do so in a continuous time environment, while we consider a discrete time environment.} This leads to a smooth response to shocks analogous to the case of a fixed probability of changing portfolios. But the latter is not analytically tractable and requires a numerical solution. We will derive a closed form solution for the exchange rate, with the exception of the case where we introduce long-term bonds to address the last puzzle.

The remainder of the paper is organized as follows. In Section 2 we discuss a two country model with short-term bonds and gradual portfolio adjustment. In Section 3 we use the solution of the model to discuss a serious of propositions related to the first four puzzles. Section 4 discusses a numerical illustration related to the first four puzzles. Section 5 extends the model to incorporate long-term bonds to address the last puzzle. Section 6 concludes.

## 2 Model with Gradual Portfolio Adjustment and Short-Term Bonds

The five puzzles can be written both in terms of real interest rates and exchange rates and in terms of nominal interest rates and exchange rates. As Engel (2016) points out, the forward discount rate puzzle applies equally when using real variables. He proceeds to develop the puzzle that high interest rate currencies are stronger than expected based on UIP by using real interest rates and exchange rates. Gali (2017) also uses real interest rates and exchange rates to develop the forward guidance exchange rate puzzle. An advantage of stating the puzzles in terms of real variables is that the real exchange rate will be stationary, while the
nominal exchange rate is generally not stationary. We therefore use real variables, although we should stress that the equations can easily be written in nominal terms as well.\footnote{In particular, when we assume identical inflation in both countries, the real interest differential is the same as the nominal interest differential and the real exchange rate is equal to the nominal exchange rate. It is well known that the real exchange rate is very closely tied to the nominal exchange rate over short horizons.}

In this section we will first describe the model and then the solution for the equilibrium real exchange rate. There are two countries, Home and Foreign, with agents that invest in one-period bonds of both countries. We adopt a simplifying overlapping generations as in Bacchetta and van Wincoop (2007, 2010). But the model differs in that we adopt a cost of changing the portfolio share instead of a fixed interval of changing portfolios, we focus on real variables and we consider investors from both countries rather than just the Home country.

### 2.1 Model Description

There are overlapping generations of agents that live two periods and are born with a wealth of 1 in real terms. Agents in the Home country born at time $t$ maximize

$$E_t \frac{C_{t+1}^{1-\gamma}}{1-\gamma} - 0.5\psi (z_t - z_{t-1})^2$$

where the second term is an adjustment cost. It captures a utility cost of choosing a different portfolio share $z_t$ invested in Foreign bonds than that of the “parents” one period ago. Consumption is equal to the portfolio return:

$$C_{t+1} = R_{t+1}^p = \left[ z_t \frac{S_{t+1}}{S_t} e^{i_t} e^{-\tau} + (1 - z_t) e^{i_t} \right] \frac{P_t}{P_{t+1}} + T_{t+1}$$

$i_t$ and $i_t^*$ are the nominal interest rate on Home and Foreign bonds. $S_t$ is the level of the nominal exchange rate, measured in terms of the Home currency per unit of the Foreign currency. $P_t$ is the price level. We will assume that inflation over the next period is known, so that $P_{t+1}$ is known. This captures the fact that there is much more uncertainty about exchange rates than inflation over the near future. We also allow for a cost $\tau$ of investing in the Foreign bonds, which is an international financial friction. The aggregate of this cost across agents is reimbursed through
$T_{t+1}$. $R^p_{t+1}$ is therefore the same in equilibrium as it would be when $\tau = 0$, but agents take $T_{t+1}$ as given, not under their control through portfolio choice.\footnote{See Bacchetta and van Wincoop (2017) and Davis and van Wincoop (2017) for the same approach.}

Define the gross real interest rates as

\[
R_t = e^{i_t} \frac{P_t}{P_{t+1}}
\]
\[
R^*_t = e^{i^*_t} \frac{P^*_t}{P^*_{t+1}}
\]

where $P^*_t$ is the Foreign price level. The real exchange rate is defined as

\[
Q_t = \frac{S_t P^*_t}{P_t}
\]

The first-order condition is then

\[
E_t e^{-\tau r^p_{t+1} + q_{t+1} - q_t + r^*_t - \tau} - E_t e^{-\tau r^p_{t+1} + r_t} - \psi(z_t - z_{t-1}) = 0
\]  
(3)

where lower case letters denote logs.

The first-order approximation of the log portfolio return is

\[
r^p_{t+1} = z_t e^{r_{t+1}} + r_t
\]

where the excess return is

\[
er_{t+1} = q_{t+1} - q_t + r^*_t - r_t
\]  
(4)

Substituting the first-order approximation of the log portfolio return into the first-order condition, assuming log-normality, and using the approximation $e^x = 1 + x$, we have

\[
E_t e^{r_{t+1} - \tau} + 0.5 \text{var}(er_{t+1}) - \gamma z_t \text{var}(er_{t+1}) - \psi(z_t - z_{t-1}) = 0
\]  
(5)

Define the steady state fraction invested in Foreign bonds as

\[
\bar{z} = \frac{0.5 \tau}{\gamma \text{var}(er_{t+1})}
\]  
(6)

We take $\bar{z}$, which is related to $\tau$, as a parameter that is given.
The optimal portfolio can then be written as

$$z_t = \frac{\psi}{\psi + \gamma \sigma^2} z_{t-1} + \frac{\gamma \sigma^2}{\psi + \gamma \sigma^2} z^f_t$$

(7)

where \(\sigma^2 = \text{var}(er_{t+1})\) and

$$z^f_t = \bar{z} + \frac{E_t er_{t+1}}{\gamma \sigma^2}$$

is the frictionless optimal portfolio in the absence of adjustment costs. (7) shows that the portfolio share \(z_t\) is a weighted average of the previous period’s portfolio share and the frictionless optimal portfolio share. The portfolio share therefore gradually adjusts to the frictionless optimal portfolio share. However, in equilibrium \(z^f_t\) will not be a fixed target as the expected excess return is endogenous and changes over time.

There is an analogous solution for the Foreign country. The real portfolio return is

$$R^{n*}_{t+1} = z^*_t R^*_t + (1 - z^*_t) e^{-\tau} R_t \frac{Q_t}{Q_{t+1}} + T^*_t$$

(8)

where \(z^*_t\) is the fraction invested in Foreign bonds by Foreign investors and \(T^*_t\) is a reimbursement of the cost of investment abroad born by the aggregate of Foreign investors. Following the same steps as for the Home country, the equation analogous to (5) is

$$E_t er_{t+1} + \tau - 0.5 \text{var}(er_{t+1}) + \gamma (1 - z^*_t) \text{var}(er_{t+1}) - \psi (z^*_t - z^*_{t-1}) = 0$$

(9)

It is easy to see that the steady state portfolio share is \(\bar{z}^* = 1 - \bar{z}\). Taking the average of (5) and (9), we have

$$E_t er_{t+1} + \gamma \sigma^2 (0.5 - z^A_t) - \psi (z^A_t - z^A_{t-1}) = 0$$

(10)

where \(z^A_t = 0.5(z_t + z^*_t)\) is the average portfolio share invested in the Foreign bond.

Next consider bond market equilibrium. It is sufficient to focus on the Foreign bond market equilibrium as by Walras’ Law the Home bond market equilibrium will then hold automatically as well. The Foreign bond market equilibrium condition is

$$z_t + z^*_t Q_t = Q_t$$

(11)

The supply and demand for Foreign bonds are expressed in real terms from the perspective of the Home country. We assume that the supply of bonds is fixed.
at 1 in real terms in terms of the purchasing power of the respective countries. The Foreign real bond supply of 1 is therefore equal to $Q_t$ in real terms from the perspective of Home agents. Similarly, the given real wealth of 1 of Foreign investors equals $Q_t$ in terms of Home purchasing power. We linearize the market clearing condition around the log of the real exchange rate $q = 0$ and portfolio shares $\bar{z}$ and $\bar{z}^*$. This gives

$$z_t^A = 0.5 + 0.5(1 - \bar{z}^*)q_t$$

(12)

It is useful to introduce a home bias parameter $h$. Home bias is usually defined as one minus the ratio of the share invested abroad and the share of the foreign asset in the world asset supply. The home bias parameter in steady state in our model is therefore

$$h = 1 - \frac{1 - \bar{z}^*}{0.5}$$

Using this, we can also write (48) as

$$z_t^A = 0.5 + \frac{1}{4}(1 - h)q_t$$

(13)

Substituting this into (10) and using the expression (4) for the excess return, we have

$$E_t q_{t+1} - \theta q_t + b\psi q_{t-1} + r_t^D = 0$$

(14)

where $r_t^D = r_t^* - r_t$ is the Foreign minus Home real interest rate differential and

$$b = \frac{1}{4}(1 - h)$$

$$\theta = 1 + \psi b + \gamma \sigma^2 b$$

$b$ is a parameter between 0 and 0.25 and $\theta > 1$.

### 2.2 Solution Real Exchange Rate

Using (14), which is a second-order difference equation in $q_t$, we can solve for the equilibrium $q_t$ as a function of the lagged real exchange rate and expected future interest rate differentials. We will treat the interest differential $r_t^D$ as given, for example driven by monetary policy shocks. Using standard solution techniques for second-order stochastic difference equations, we have

$$q_t = \alpha q_{t-1} + E_t \sum_{i=0}^{\infty} \frac{1}{D^{i+1}} r_t^{D+i}$$

(15)
where $\alpha$ and $D$ are the roots of the characteristic equation of (14):

$$\alpha = \frac{\theta - \sqrt{\theta^2 - 4\psi b}}{2}$$  \hspace{1cm} (16)$$

$$D = \frac{\theta + \sqrt{\theta^2 - 4\psi b}}{2}$$  \hspace{1cm} (17)$$

It is easily verified that $0 \leq \alpha < 1$ and $D > 1$. The equilibrium real exchange rate therefore depends on the lagged real exchange rate and a present discounted value of expected future real interest rate differentials. Since $D > 1$ it is immediate that expected real interest rates in the more distant future have a smaller effect on the equilibrium real exchange rate than expected real interest rates in the near future. This addresses the forward guidance exchange rate puzzle. We will explore this more, and develop the intuition behind it, in the next two sections.

We will focus on the case where the real interest differential follows a simple AR(1) process:

$$r_{t}^{D} = \rho r_{t-1}^{D} + \varepsilon_{t}$$  \hspace{1cm} (18)$$

In that case (15) gives us

$$q_{t} = \alpha q_{t-1} + \frac{1}{D - \rho} r_{t}^{D}$$  \hspace{1cm} (19)$$

We can also write this as a function of current and past real interest differentials:

$$q_{t} = \frac{1}{D - \rho} \sum_{i=0}^{\infty} \alpha^{i} r_{t-i}^{D}$$  \hspace{1cm} (20)$$

or as a function of current and past real interest rate shocks:

$$q_{t} = \frac{1}{D - \rho} \sum_{i=0}^{\infty} \nu_{i} \varepsilon_{t-i}$$  \hspace{1cm} (21)$$

where

$$\nu_{i} = \begin{cases} 
\frac{\alpha^{i+1} - \rho^{i+1}}{\alpha - \rho} & \text{if } \alpha \neq \rho \\
(i+1)\rho^{i} & \text{if } \alpha = \rho 
\end{cases}$$  \hspace{1cm} (22)$$

A couple of comments about the parameters $\alpha$ and $D$ are in order as they are key to the solution. Appendix A derives the following Lemma:

**Lemma 1** The following properties describe the relationship between $\alpha$, $D$ and the portfolio adjustment cost parameter $\psi$:
• As $\psi$ rises from 0 to $\infty$, $\alpha$ rises monotonically from 0 to 1
• As $\psi$ rises from 0 to $\infty$, $D$ rises monotonically from $1 + \gamma \sigma^2$ to $\infty$

Higher portfolio adjustment costs imply that the real exchange rate depends to a greater extent on the value of the real exchange rate during the last period and future expected real interest rates are discounted more in the equilibrium real exchange rate.

2.3 Excess Return predictability Coefficients

Consider the following regression:

$$er_{t+k} = \alpha + \beta_k r^D_t + \varepsilon_{t+k}^er$$  \hspace{1cm} (23)

Several of the puzzles are related to the excess return predictability coefficients $\beta_k$, which tells us the effect of the current real interest differential on the expected excess return $k$ periods from now. The forward discount puzzle focuses on $k = 1$, with one period usually being a month or a quarter. We therefore consider the predictability of the excess return over the next month or quarter. For the predictability reversal puzzle and the Engel puzzle we are also interested in the $\beta_k$ for $k > 1$, which relates to the effect of the current interest differential on the excess return further into the future.

In the model, the value of $\beta_k$ is equal to

$$\beta_k = \frac{\text{cov}(er_{t+k}, r^D_t)}{\text{var}(r^D_t)}$$  \hspace{1cm} (24)

Using the solution for the real exchange rate under the assumed AR(1) process for the real interest differential, Appendix B shows that this can be written as

$$\beta_k = \begin{cases} 
\lambda_1 \rho^{k-1} + \lambda_2 \alpha^{k-1} & \text{if } \alpha \neq \rho \\
\frac{\rho^{k-1}}{D - \rho} \left( D - \frac{1}{1 + \rho} - (1 - \rho)(k - 1) \right) & \text{if } \alpha = \rho 
\end{cases}$$  \hspace{1cm} (25)

where

$$\lambda_1 = \frac{1}{D - \rho} \left( D - \rho \frac{\alpha - 1}{\alpha - \rho} \right)$$  \hspace{1cm} (26)

$$\lambda_2 = \frac{\alpha - 1}{D - \rho} \left[ \frac{\rho}{\alpha - \rho} + \frac{1}{1 - \alpha \rho} \right]$$  \hspace{1cm} (27)
Although $\beta_k$ is a continuous function of $\alpha$ (and therefore of $\psi$), $\lambda_1$ and $\lambda_2$ are not defined at $\alpha = \rho$, which is why the expression for $\beta_k$ at $\alpha = \rho$ is reported separately.

Define

$$\tilde{\psi}_1 = \frac{\rho}{1 - \rho} \gamma \sigma^2$$

(28)

$$\tilde{\psi}_2 = \frac{\rho}{1 - \rho} \gamma \sigma^2 + \frac{\rho}{b}$$

(29)

Clearly $\tilde{\psi}_2 > \tilde{\psi}_1$. $\tilde{\psi}_1$ corresponds to the value of $\psi$ where $\lambda_1 = 0$ (see Appendix C), while $\tilde{\psi}_2$ corresponds to the value of $\psi$ where $\alpha = \rho$. In Appendix C we prove the following Lemma:

**Lemma 2** There are three regions that determine the sign of $\lambda_1$ and $\lambda_2$:

- $0 < \psi < \tilde{\psi}_1$: $\lambda_1 > 0$ and $\lambda_2 > 0$
- $\tilde{\psi}_1 < \psi < \tilde{\psi}_2$: $\lambda_1 < 0$ and $\lambda_2 > 0$
- $\psi > \tilde{\psi}_2$: $\lambda_1 > 0$ and $\lambda_2 < 0$

When $\psi = 0$, $\lambda_1 > 0$ and $\lambda_2 = 0$. When $\psi = \tilde{\psi}_1$, $\lambda_1 = 0$ and $\lambda_2 > 0$.

### 3 Some Propositions Related to Puzzles 1-5

We will now use the simple model discussed above to address the first four puzzles. In this section we do so by discussing a series of propositions related to these puzzles. In the next section we consider a numerical illustration.

#### 3.1 Delayed Overshooting Puzzle

Define

$$\bar{t} = \begin{cases} 
\frac{\ln(1 - \rho) - \ln(1 - \alpha)}{\ln(\alpha) - \ln(\rho)} & \text{if } \alpha \neq \rho \\
\frac{\rho}{1 - \rho} & \text{if } \alpha = \rho
\end{cases}$$

(30)

Using (21), Appendix D proves the following Proposition:

**Proposition 1** Consider the impulse response of the real exchange rate to a positive interest rate shock
• if $\alpha + \rho < 1$: the real exchange rate appreciates at the time of the shock and subsequently gradually depreciates back to the steady state.

• if $\alpha + \rho > 1$: there is delayed overshooting. The real exchange rate appreciates at the time of the shock, then continues to appreciate, and then, starting at time $t > \tilde{t} > 1$ starts to gradually depreciate back to the steady state.

Since Lemma 1 tells us that $\alpha$ rises from 0 to 1 as we raise the gradual portfolio adjustment parameter $\psi$, Proposition 1 implies that for sufficiently large $\psi$, and assuming $\rho > 0$, there is delayed overshooting of the type reported by Eichenbaum and Evans (1995) and others. They show that that after monetary policy tightening the currency continues to appreciate for another 8-12 quarters before it starts to depreciate. When the extent of gradual adjustment is less, such that $\alpha < 1 - \rho$, there is no delayed overshooting. To understand the intuition, consider an increase in the dollar interest rate. There will be an immediate appreciate of the dollar as investors shift to dollar bonds. Subsequent to the shock, there are two opposing forces at work. On the one hand, the the dollar interest rate starts to gradually decline again, which leads to a shift away from dollars and therefore a gradual depreciation. On the other hand, to the extent that portfolios are slow to adjust, there will be a continued flow towards dollar bonds, which leads to a continued appreciation. When $\psi$ is sufficiently large, the second force dominates and there will be delayed overshooting.

(30) tells us for how long the real appreciation will last in the case of delayed overshooting. Appendix D shows that the derivative of $\tilde{t}$ with respect to $\alpha$ is positive. A larger gradual portfolio adjustment parameter $\psi$, which raises $\alpha$ (Lemma 1), will then lead to a longer duration of the delayed overshooting. In the extreme where $\alpha$ approaches 1, $\tilde{t}$ approaches infinity.

When $\alpha < 1 - \rho$, the real exchange rate starts to immediately depreciate subsequent to the initial appreciation at the time of the shock. But the speed of depreciation does critically depends on $\alpha$, and therefore the gradual portfolio adjustment parameter $\psi$. In particular, if the shock occurs at $t = 0$, leading to a real appreciation at time zero, we have $q_1 - q_0 = \alpha + \rho - 1$. Therefore, the larger $\psi$ and $\alpha$, the less the real exchange rate will appreciate. When $\alpha$ is close to $1 - \rho$, the real exchange rate depreciation will be close to zero. This feature is important to some of the other puzzles and delayed overshooting is indeed not critical to addressing these other puzzles. This is important as some of
the literature subsequent to Eichenbaum and Evans (1995) and Grilli and Roubini (1996) has shown that the puzzle does not apply to all time periods and countries.

While Proposition 1 emphasizes the behavior of the real exchange rate subsequent to the initial shock, there is a close relationship between the initial response to the shock and the subsequent response. (21) implies that in the immediate response to a shock at time zero we have \( \Delta q_0 = \left( \frac{1}{D - \rho} \right) \epsilon_0 \). We know from Lemma 1 that \( D \) is a positive function of \( \psi \), so that a higher portfolio adjustment cost implies a smaller initial response of the real exchange rate. As \( \psi \) rises from 0 to infinity, the initial appreciation becomes smaller, while the subsequent response changes from subsequent depreciation \( (\psi = 0) \) to weaker subsequent depreciation \( (0 < \alpha < 1 - \rho) \) to subsequent appreciation \( (\alpha > 1 - \rho) \).

Figure 1 provides a numerical illustration. The chart on the left shows the impulse response of the real exchange rate when \( \psi = 12 \) and \( \gamma = 10 \). We refer to this as the benchmark case. The chart on the right shows the time to maximum overshooting for \( \psi \) varying from 0 to 30 and \( \gamma \) taking on the values 1, 10 and 50. In both cases we set \( h \) equal to 0.66. This is the average for the 6 countries in Engel (2016), plus the United States, during Q2, 2017. We set \( \rho = 0.8 \) and \( \sigma = 0.057 \), using quarterly data in Bacchetta and van Wincoop (2010).

Chart A of Figure 1 shows that the real exchange rate overshoots, reaching a maximum after 9 quarters. Eichenbaum and Evans (1995) find that the maximum exchange rate impact occurs on average at 35 months, or 11.5 quarters, for the 5 currencies that they consider. This is not too far from the 9 quarters in Figure 1. Chart B shows that with the exception of values of \( \psi \) very close to zero, the model implies delayed overshooting. Consistent with Proposition 1, the time to maximum impact rises significantly with \( \psi \). It is also larger the lower the rate of risk-aversion.

### 3.2 Forward Discount Puzzle

In Appendix E we use (25) to prove the following proposition:

**Proposition 2** The Fama predictability coefficient \( \beta_1 \) is positive, and larger when there is gradual portfolio adjustment \( (\psi > 0) \)

---

5The countries other than the US are Canada, France, Germany, Italy, Japan and UK. We combine BIS data on debt securities outstanding with external assets and liabilities for debt securities from the IMF International Investment Position Statistics.
Proposition 2 says that $\beta_1 > 0$, so that a positive expected excess return is expected on the high interest rate currency, consistent with the forward discount puzzle. It furthermore says that $\beta_1$ is larger when we introduce a cost of adjusting portfolios ($\psi > 0$). Even without this cost, there is a little bit of excess return predictability in the model through a risk premium channel. Specifically, a higher Foreign real interest rate will lead to a real appreciation of the Foreign currency, which increases the value of the Foreign bond supply in real terms from the perspective of Home agents. Investors then demand a positive expected excess return on the Foreign bond. We will see in the next section that this standard risk premium channel is quantitatively weak. Proposition 2 says that predictability is strengthened by introducing gradual portfolio adjustment.

Proposition 1 on delayed overshooting is a useful starting point to understand the role of gradual portfolio adjustment in accounting for the forward discount puzzle. When $\alpha + \rho > 1$, so that there is delayed overshooting, the real exchange rate of the high interest rate currency is expected to appreciate for at least one more period after the initial appreciation. The higher interest rate currency will then have a positive expected excess return both due to the higher interest rate and the expected appreciation. When $\psi = 0$, the real exchange rate is expected to depreciate subsequent to the shock, reducing the expected excess return. So the gradual adjustment, by generating a gradual portfolio shift to the high interest rate currency and therefore a gradual appreciation, increases the Fama predictability coefficient $\beta_1$.

But delayed overshooting is not critical to Proposition 2. Even when $\alpha + \rho < 1$, so that there is not delayed overshooting, gradual portfolio adjustment leads to a higher Fama coefficient $\beta_1$. We have seen that the rate of depreciation subsequent to the shock is smaller due to gradual portfolio adjustment. The subsequent depreciation is caused by a decline over time in the interest rate differential, leading to a portfolio shift away from the currency whose interest rate is falling. But this is partially offset by the gradual portfolio shift to the higher interest rate currency. Therefore the larger $\psi$, the weaker the depreciation after the shock, to the point where the depreciation in the period after the shock is close to zero when $\alpha$ is close to $1 - \rho$. The weaker subsequent depreciation implies a higher expected excess return on the high interest rate currency and therefore a larger Fama coefficient $\beta_1$.

In the benchmark case of Figure 1, where $\psi = 12$ and $\gamma = 10$, the excess
return predictability coefficient $\beta_1$ is equal to 3.33. Figure 2 shows how $\beta_1$ varies with $\psi$ and $\gamma$. It rises until $\psi$ is about 12 and then gradually drops. When $\psi$ is very large the exchange rate is very slow to respond due to the weak portfolio response, diminishing the excess return predictability associated with the Foreign currency appreciation after the shock. Figure 2 also shows that the excess return predictability coefficient $\beta_1$ is larger when risk-aversion $\gamma$ is smaller. A very large $\gamma$ again leads to a weak portfolio response, diminishing excess return predictability associated with the Foreign currency appreciation.

3.3 Predictability Reversal Puzzle

In Appendix F we prove the following proposition:

**Proposition 3** The following holds for $\beta_k$:

- if $\psi \leq \bar{\psi}_1$: $\beta_k$ is positive for all $k$ and drops monotonically to zero as $k \to \infty$

- if $\psi > \bar{\psi}_1$: there is a $\bar{k} > 1$ such that $\beta_k$ is positive for $k < \bar{k}$ and negative for $k \geq \bar{k}$. It converges to zero as $k \to \infty$.

Proposition 3 says that when the gradual adjustment parameter is low ($\psi \leq \bar{\psi}_1$), currencies whose current interest rate is relatively high continue to have positive expected excess returns in all future periods, although the predictability $\beta_k$ vanishes to zero over time. But when the gradual adjustment parameter is sufficiently high ($\psi > \bar{\psi}_1$), there will be a predictability reversal. While currencies whose interest rate is currently high are initially expected to have a positive expected excess return, after a certain period of time they are expected to have negative expected excess return. Bacchetta and van Wincoop (2010) first documented this reversal of the sign of predictability for nominal interest rates and exchange rates. They find that a high interest rate currency has a positive expected excess return for about 5-10 quarters, after which it has a negative expected excess return. Engel (2016) reports the same finding for real interest rates and exchange rates.

We should first point out that delayed overshooting is not critical to the predictability reversal result. To see this, delayed overshooting happens when $\alpha > 1 - \rho$, which corresponds to

$$\psi > \bar{\psi} = \frac{1 - \rho}{b} + \frac{1 - \rho}{\rho} \gamma \sigma^2$$
There can be predictability reversal without delayed overshooting ($\bar{\psi}_1 < \psi < \bar{\psi}$) when

$$\rho(1 - \rho)^2 > (2\rho - 1)\gamma\sigma^2b$$

As we will see in the next section, this condition can easily be satisfied for plausible parameters.

Nonetheless, to understand the predictability reversal result, it is useful to start with the case of delayed overshooting. Ultimately, when $t > \bar{t}$, the real exchange rate will start to depreciate. $\bar{t}$ may be a large number, well into the future. By that time the interest differential will be close to zero. The excess return is then largely driven by the exchange rate. A depreciating currency then has a negative excess return. Even in the absence of delayed overshooting, we have seen that the depreciation subsequent to the shock may be initially weak, delaying the ultimate return of the real exchange rate to its steady state. Since the interest differential is declining to zero, this can after a certain period lead to negative excess returns.

Engel (2016) claims that models with gradual portfolio adjustment cannot account for the predictability reversal. To show this, Engel (2016) does not develop a model of gradual portfolio adjustment, but instead considers a reduced form exchange rate equation that he believes to be implied by models of portfolio adjustment. This equation involves an AR(1) process for $q_t - q_t^{IP}$, where $q_t^{IP}$ is the exchange rate under interest rate parity where the expected excess return is zero:

$$E_t(q_t^{IP} + 1 - q_t^{IP} + r_t^* - r_t) = 0.$$  

The exchange rate under interest rate parity is

$$q_t^{IP} = \sum_{i=0}^{\infty} E_t r_{t+i}^{IP}$$  \hspace{1cm} (31)$$

Engel conjectures that under gradual portfolio adjustment the real exchange rate gradually approaches the level that would apply under interest rate parity. This is not the case though. First, as can be seen from (14), the real exchange rate is driven by an AR(2) process, not an AR(1) process. Second, as we will see more clearly in a numerical illustration in the next section, the gap between $q_t$ and $q_t^{IP}$ will generally grow significantly over time after the shock.

Chart A of Figure 3 reports $\beta_k$ for $k$ from 1 to 50 for the bechmark case where $\psi = 12$ and $\gamma = 10$. The reversal of the predictability coefficient from positive to negative occurs after 7 quarters. This is not too far from the reversal after 5-10 quarters reported in Bacchetta and van Wincoop (2010). Engel (2016) reports a reversal after about one year. The charts reported in Engel (2016) for $\beta_k$ for 6
currencies and the G6 are very similar to the benchmark case in chart A of Figure 3. Chart B considers the impact of $\psi$ and $\gamma$ on the time $k$ where $\beta_k$ reverses sign from positive to negative. This rises with $\psi$ and falls with $\gamma$, consistent with the longer time to maximum impact of the exchange rate for higher $\psi$ and lower $\gamma$ shown in Chart 1B.

### 3.4 Engel Puzzle

The Engel puzzle says that high interest rate currencies tend to be strong relative to the interest parity exchange rate. More formally:

$$\text{cov}(q_t - q_t^{IP}, r_t^D) > 0$$  \hspace{1cm} (32)

Engel (2016) provides evidence that this condition holds in the data for 6 currencies. We will refer to it as the Engel condition. Using that $\sum_{k=0}^{\infty}(q_{t+k+1} - q_t + r_t^{D, k}) = q_\infty - q_t + \sum_{k=0}^{\infty}r_t^{D, k}$, we have

$$\text{var}(r_t^D)\sum_{k=1}^{\infty}\beta_k = \text{cov}\left(q_\infty - q_t + \sum_{k=0}^{\infty}r_t^{D, k}, r_t^D\right) = \text{cov}(q_t^{IP} - q_t, r_t^D)$$  \hspace{1cm} (33)

The last equality uses that $q_\infty$ and $\sum_{k=0}^{\infty}(r_t^{D, k} - E_t r_t^{D, k})$ are unaffected by shocks that affect $r_t^D$. It follows that the Engel condition can be written as

$$\sum_{k=1}^{\infty}\beta_k < 0$$  \hspace{1cm} (34)

which is an equivalency used by Engel (2016) as well. Predictability reversal is a necessary condition for this to hold, so that $\psi > \bar{\psi}_1$ is necessary, but it is not a sufficient condition. For the currency whose real interest rate is currently relatively high, the negative predictability of excess returns when $k \geq \bar{k}$ must more than offset the positive expected excess returns when $k < \bar{k}$.

Define $\bar{\psi}_1$ and $\bar{\psi}_2$ as positive values of $\psi$, with $\bar{\psi}_1 < \bar{\psi}_2$, where $\text{cov}(q_t - q_t^{IP}, r_t^D) = 0$. Appendix G describes these values and proves the following proposition:

**Proposition 4** Necessary and sufficient conditions for the Engel condition to hold are

1. $\gamma \sigma^2 b < \frac{1-\rho}{\rho} \left(1 - \sqrt{1-\rho}\right)^2$
Proposition 4 imposes several restrictions on parameters in order for the Engel condition to be satisfied. First, risk or risk-aversion cannot be too large. Second, the Engel condition is only satisfied for intermediate values of the interest rate persistence \( \rho \). It will not hold when \( \rho \) is very close to either 0 or 1. Third, the gradual portfolio adjustment parameter \( \psi \) cannot be too large or too small. While one might think that these restrictions imply that the Engel condition is not likely to be satisfied, we will see in the numerical exercise of the next section that it will be satisfied for a very broad range of parameters.

In order to understand the intuition behind the proposition, and the role of the various parameters, we can write the Engel condition as

\[
\text{cov}(q_t, r^D_t) > \frac{1}{1 - \rho} \text{var}(\varepsilon_t)
\]  

(35)

In what follows assume that the interest rate of the Foreign currency rises, so that \( r^D_t \) rises and the Foreign currency appreciates (\( q_t \) rises). The right hand side of (35) captures the fact that interest differentials themselves (ignoring exchange rate changes) lead to positive expected excess returns on Foreign currency. This goes against the Engel condition, which says that the sum of all future excess returns is negative on the high interest rate currency. The left hand side says that the more the Foreign currency appreciates, the more likely the Engel condition is satisfied. A large appreciation implies that subsequent depreciations must be large, which lowers future excess returns.

First consider the role of the gradual portfolio adjustment parameter \( \psi \). We already knew from Proposition 3 that it cannot be too small as \( \psi > \bar{\psi}_1 \) is needed to obtain predictability reversal. At the same time, when \( \psi \) gets too large, the predictability reversal is not strong enough to satisfy the Engel condition. The gradual portfolio adjustment parameter only affects the left hand side of (35). When \( \psi \) is very large, portfolios are very slow to respond, so that \( q_t \) rises little at the time of the shock and the initial periods after the shock when the interest differential is large. (35) will then not be satisfied. While the real exchange rate will eventually depreciate, the depreciation will be small as the initial appreciation is small.

Next consider the role of risk or risk-aversion. If \( \gamma \sigma^2 \) is extremely large, portfolios are not very responsive to expected returns and therefore the Foreign currency
appreciates very little. Just like for a very high $\psi$, this implies that the left hand side of (35) is insufficiently large for the Engel condition to be satisfied. The future depreciations will be weak as the initial appreciation is weak. Interest differentials are then the main driver of future excess returns.

Finally, consider the persistence $\rho$ of the real interest rate. The right hand side of (35) goes to infinity when $\rho$ approaches 1, so that the Engel condition will not be satisfied. If the interest differential is very persistent, the Foreign currency will continue to experience high interest rates for a very long time, which by itself causes positive excess returns for a long time. But Proposition 4 tells us that the Engel condition is also not satisfied when $\rho$ is very close to zero, so that the shock has virtually no persistence. Such lack of persistence implies a weak appreciation of the real exchange rate at the time of the shock. The left hand side of (35) will then be small. Subsequent depreciation of the Foreign currency will then not be sufficient to satisfy the Engel condition.

Figure 4 shows that the Engel result holds for a very wide range of values of $\psi$, from just above zero to close to 20. Consistent with Proposition 4, the Engel result is stronger the lower the rate of risk-aversion $\gamma$ and peaks for an intermediate value of $\psi$.

3.5 Forward Guidance Exchange Rate Puzzle

The following proposition addresses the forward guidance puzzle posed by Gali (2017):

**Proposition 5** The real exchange rate $q_t$ gives less weight to expected interest differentials in the distant future than the near future. The higher the gradual portfolio adjustment parameter $\psi$, the more future expected interest differentials are discounted in the equilibrium real exchange rate.

Proposition 5 follows directly from (15) and Lemma 1. Future expected interest differentials are discounted at the rate $D$, which is larger than 1 and rises with $\psi$.

The puzzle develops if one linearizes the first-order condition for portfolio choice. Linearization of portfolio Euler equations leads to an approximation that abstracts from risk, such that the expected excess return is always zero. This leads
to what Engel (2016) calls the interest parity real exchange rate:

$$q_t^{IP} = \sum_{i=0}^{\infty} E_t r_{t+i}^D$$

In this interest parity real exchange rate there is no discounting. Expected interest differentials in the near future have the same weight as expected interest differentials in the distant future.

When $\psi = 0$, $D = 1 + \gamma \sigma^2 b$, so that there will already be some discounting. This occurs because non-zero risk premia are needed to induce investors to buy the assets. Consider an increase in the expected Foreign interest rate next period, at time 1. This leads to an appreciation of the Foreign currency today, at time 0 ($q_0$ rises). But the higher expected Foreign interest rate at time 1 is discounted twice. First, the appreciation of the Foreign currency at time 1 raises the share of Foreign bonds in the total asset supply to $0.5 + bq_1$, which requires a positive expected excess return on Foreign bonds. This risk premium on Foreign bonds mitigates the appreciation of the Foreign currency at time 1. Since $q_0$ depends on the expectation of $q_1$, this lowers the appreciation at time 0. Second, the appreciation at time 0 leads to a higher relative Foreign bond supply at time 0, which implies a period 0 risk premium on the Foreign bond that further diminishes the time 0 appreciation. The further into the future the Foreign interest rate is expected to rise, the more it is discounted in the present due to higher risk premia in all future periods.

In practice though, as we will see in the next section, the discount rate $D$ is not very large when $\psi = 0$. Proposition 5 says that the discount rate rises as we increase the gradual portfolio adjustment parameter $\psi$. To see this, consider again an increase in the expected Foreign interest rate at time 1. A positive $\psi$ leads to further discounting. First, $q_1$ is not expected to rise as much when $\psi > 0$. Gradual portfolio adjustment implies a smaller portfolio shift to the Foreign currency at time 1, which leads to a smaller appreciation at time 1. This lowers the expected excess return on the Foreign bond at time 0 and therefore leads to a smaller appreciation at time 0 as well. Second, a given expected excess return on the Foreign currency at time zero implies a smaller portfolio shift to the Foreign bond at time 0 under gradual portfolio adjustment, which diminishes the appreciation of the Foreign currency at time 0. If we continue this argument for expected changes in the Foreign interest rate even further into the future, the impact on the time 0
real exchange rate will be further discounted.

The quarterly discount rate $D$ in our benchmark numerical example with $\psi = 12$ and $\gamma = 10$ is 1.07. This corresponds to a 31 percent annual discount rate. Future expected interest rates are therefore heavily discounted in the example. This is consistent with results reported by Gali (2017), which imply that expected interest rates more than two years into the future have an effect on the current real exchange rate that is very small compared to the impact of expected interest rates over the next two years.\footnote{Gali (2017) regresses $q_t$ on $\sum_{i=0}^{23} E_t r^D_{t+i}$ and $\sum_{i=24}^{\infty} E_t r^D_{t+i}$. We cannot do so in our model as both are proportional to $r^D_t$ and therefore collinear. They would no longer be collinear if we adopted an AR(2) process. More generally, the precise coefficients that we would obtain for a Gali type regression depend on what we assume about the information about future expected interest differentials, which is auxiliary to the gradual portfolio adjustment aspect of the model.}

For comparison, when $\psi = 0$ (holding all other parameters the same), the discount rate is $D = 1.0028$, which is only 1.1 percent at an annual basis. In that case expected interest rates ten years into the future have an effect on the current exchange rate that is only 11 percent less than effect of current interest rates.

4 LSV Puzzle: Model with Long Term Bonds

In order to address the last puzzle we need to extend the model by introducing long-term bonds. In that case we need to solve not only for the equilibrium real exchange rate, but also long-term bond prices in both countries. In this extension an analytical solution is no longer feasible. We will describe the extended model, leaving all algebraic details to a separate Online Appendix.

In addition to the one-period bonds of the model in Section 2, there are now also long-term bonds in both countries. This makes for a total of 4 assets. Agents in the Home country maximize

$$E_t \frac{C_{t+1}^{1-\gamma}}{1-\gamma} - \frac{1}{4} \psi \sum_{i=1}^{4} (z_{it} - z_{i,t-1})^2$$

(36)

$z_{1t}$ is the fraction of wealth invested in Foreign short term bonds. $z_{2t}$ is the fraction invested in Foreign long term bonds and $z_{3t}$ is the fraction invested in Home long term bonds. The remaining fraction $z_{4t} = 1 - z_{1t} - z_{2t} - z_{3t}$ is invested in Home short term bonds. The adjustment cost term in (36) is the same as in (1) when...
we set the long term bond portfolio shares equal to 0. In that case $z_{1t} = z_t$ and $z_{4t} = 1 - z_t$.

Let $R_{t+1}^L$ be the real return on Home long term bonds from the perspective of Home agents and $R_{t+1}^{L,*}$ the real return on Foreign long term bonds from the perspective of Foreign agents. The gross real interest rates on one-period bonds will continue to be denoted as $R_t$ and $R_t^*$. Consumption of Home agents is equal to the portfolio return:

$$C_{t+1} = R_t + z_{1t} \left( \frac{Q_{t+1}}{Q_t} R_t e^{-\tau} - R_t \right) + z_{2t} \left( \frac{Q_{t+1}}{Q_t} R_{t+1}^{L,*} e^{-\tau L} - R_t \right) + z_{3t} \left( R_{t+1}^L - R_t \right) + T_{t+1}$$

(37)

As before, the cost of investing abroad is $\tau$ for short-term bonds. It is $\tau_L$ for long term bonds. The aggregate of these costs are reimbursed through $T_{t+1}$.

Long term bonds in both countries earn real coupons of $\kappa$, $(1 - \delta)\kappa$, $(1 - \delta)^2\kappa$, and so on. A smaller $\delta$ implies a longer maturity of debt. The real returns on Home and Foreign long term bonds, from the perspective of respectively Home and Foreign agents, are then

$$R_{t+1}^L = \frac{(1 - \delta)P_{t+1}^L + \kappa}{P_t^L}$$

(38)

$$R_{t+1}^{L,*} = \frac{(1 - \delta)P_{t+1}^{L,*} + \kappa}{P_t^{L,*}}$$

(39)

Here $P_t^L$ and $P_t^{L,*}$ are the prices of newly issued bonds at time $t$, measured in real terms from the perspective of Home and Foreign agents.

Log excess returns are defined as the log real asset returns from the perspective of Home agents minus the real interest rate $r_t$ of the Home country. The vector of excess returns on the first three assets, not including the cost of investing abroad, is

$$\text{er}_{t+1} = \begin{pmatrix} \text{er}_{1,t+1} \\ \text{er}_{2,t+1} \\ \text{er}_{3,t+1} \end{pmatrix} = \begin{pmatrix} q_{t+1} - q_t + r_t^* - r_t \\ q_{t+1} - q_t + r_{t+1}^{L,*} - r_t \\ r_{t+1}^L - r_t \end{pmatrix}$$

(40)

As before, the cost of investing abroad is $\tau$ for short-term bonds. It is $\tau_L$ for long term bonds. The aggregate of these costs are reimbursed through $T_{t+1}$.

Log excess returns are defined as the log real asset returns from the perspective of Home agents minus the real interest rate $r_t$ of the Home country. The vector of excess returns on the first three assets, not including the cost of investing abroad, is

$$\text{er}_{t+1} = \begin{pmatrix} \text{er}_{1,t+1} \\ \text{er}_{2,t+1} \\ \text{er}_{3,t+1} \end{pmatrix} = \begin{pmatrix} q_{t+1} - q_t + r_t^* - r_t \\ q_{t+1} - q_t + r_{t+1}^{L,*} - r_t \\ r_{t+1}^L - r_t \end{pmatrix}$$

(40)

Define $\Sigma$ as the variance of $\text{er}_{t+1}$. Using log normality of consumption and returns, the Online Appendix shows that the first-order conditions of Home agents
can be written as
\[ E_t e_{t+1} - \begin{pmatrix} \tau \\ \tau_L \\ 0 \end{pmatrix} + 0.5 \text{diag}(\Sigma) - \gamma \Sigma z_t = \frac{\psi}{2R} (\hat{z}_t - \hat{z}_{t-1}) \]  
(41)

where \( z_t = (z_{1t}, z_{2t}, z_{3t})' \) and \( \hat{z}_t \) subtracts \( z_{4t} \) from each element of \( z_t \). \( R \) is the steady state gross real interest rate. The analogous first-order conditions for Foreign agents are

\[ E_t e_{t+1} + \begin{pmatrix} \tau \\ \tau_L \end{pmatrix} - (1 - \gamma) \Sigma_1 + 0.5 \text{diag}(\Sigma) - \gamma \Sigma z_t^* = \frac{\psi}{2R} (\hat{z}_t^* - \hat{z}_{t-1}^*) \]  
(42)

where \( z_t^* = (z_{1t}^*, z_{2t}^*, z_{3t}^*)' \) is the vector of portfolio shares of Foreign agents and \( \hat{z}_t^* \) subtracts \( z_{4t}^* \) from each element of \( z_t^* \). \( \Sigma_1 \) is the first column of \( \Sigma \).

The asset market equilibrium conditions can be written

\[ z_{1t} + Q_t z_{1t}^* = Q_t b^S \]  
(43)

\[ z_{2t} + Q_t z_{2t}^* = Q_t P_{tL}^* b_t \]  
(44)

\[ z_{3t} + Q_t z_{3t}^* = P_t^L b_t \]  
(45)

\[ z_{4t} + Q_t z_{4t}^* = b^S \]  
(46)

Here \( b^S \) is the constant supply of the short-term bond, while \( b_t \) is the supply of the long-term bond. Define \( b^L = \bar{P}L^t \bar{b} \) as the steady state real value (in local purchasing power) of the long-term bonds. We assume that \( b^S + b^L = 1 \) and in deviation from steady state

\[ b_t = -p_t^{L,A} \]  
(47)

where \( p_t^{L,A} = 0.5(p_t^L + p_t^{L,*}) \) is the average log long term bond price. This assures that when we add up all (log-linearized) market clearing conditions we get an identity, which must be the case due to Walras’ Law (the last market clearing condition is redundant). (47) is not important in what follows as excess returns depend on relative log bond prices, not average log bond prices.7

7The reason we need to have (47) is a bit technical. One can think of the bonds as issued by Home and Foreign governments. Since there is no investment in the model, it must be the case that world saving (private plus government) is zero in equilibrium. Since there is no endogenous mechanism in the model to equate world saving to zero, we assume that world government saving adjusts to make world saving equal to zero. This happens when (47) is satisfied, which implies that the average world real bond supply remains constant.
Assuming (47), linearizing the first three market clearing conditions, we have

\[ z^A_t = 0.5 \begin{pmatrix} b^S \\ b^L \\ b^L \end{pmatrix} + 0.5 \begin{pmatrix} b^S \\ b^L \\ 0 \end{pmatrix} q_t - 0.5 z^* q_t + 0.25 b^L p_{t,D}^L \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \]  

(48)

where \( z^A_t \) is the average of \( z_t \) and \( z_t^* \), \( z^* \) is the steady state of \( z_t^* \) and \( p_{t,D}^L = p_t^L - p_t^{L,*} \) is the relative log long term bond price.

By symmetry \( \bar{z}^*_1 + \bar{z}^*_4 = b^S \) and \( \bar{z}^*_2 + \bar{z}^*_3 = b^L \). We can choose \( \tau \) and \( \tau_L \) to set \( \bar{z}^*_4 \) and \( \bar{z}^*_3 \) at any value. We will assume that these values are such that they generate the same home bias \( h \) for short and long-term bonds. This happens when

\[ \bar{z}^*_1 = 0.5(1 - h)b^S \]  

(49)

\[ \bar{z}^*_3 = 0.5(1 - h)b^L \]  

(50)

We then also have \( \bar{z}^*_2 = 0.5(1 + h)b^L \) and \( \bar{z}^*_1 = 0.5(1 + h)b^S \).

After substituting the market clearing conditions (48) into the average of the Home and Foreign first-order conditions (41)-(42), we obtain a dynamic system of three equations in \( q_t \), \( p_{t,D}^L \) and \( p_{t,A}^L \) in deviation from steady states. We assume that both Home and Foreign interest rates \( r_t \) and \( r_t^* \) follows AR(1) processes with AR coefficients \( \rho \). This is therefore also the case for the average interest rate \( r_t^A = 0.5(r_t + r_t^*) \) and the interest differential \( r_t^D = r_t^* - r_t \). The Online Appendix shows that one of the three equations of the dynamic system can be used to solve for the average long term bond price:

\[ p_{t}^{L,A} = -\frac{1}{1 - \lambda \rho} r_t^A \]  

(51)

with \( \lambda = (1 - \delta)/R \). A higher average world real interest rate reduces the average long term bond price.

The two remaining equations of the dynamic system can be written as

\[ A_1 E_t \begin{pmatrix} q_{t+1} \\ p_{t+1}^{L,D} \end{pmatrix} + A_2 \begin{pmatrix} q_t \\ p_{t,D}^{L,D} \end{pmatrix} + A_3 \begin{pmatrix} q_{t-1} \\ p_{t-1}^{L,D} \end{pmatrix} + A_4 r_t^D = 0 \]  

(52)

The matrices \( A_1 \) through \( A_4 \) are described in the Online Appendix. They have coefficients that depend on model parameters as well as the variance \( \Sigma \) of excess returns. This dynamic system can be used to solve for \( (q_t, p_{t,D}^{L,D})' \):

\[ \begin{pmatrix} q_t \\ p_{t}^{L,D} \end{pmatrix} = \sum_{k=0}^{\infty} M^k_1 M_2 r_{t-k}^D \]  

(53)
where $M_1$ and $M_2$ are two by two matrices and $M_1^0$ is the identity matrix.

In applying the model to the data, we need to make assumptions about the parameters $h$, $\gamma$, $\psi$, $\rho$, $\delta$, $R$ and the variance $\Sigma$ of excess returns. As in Section 3, we assume $h = 0.66$, $\rho = 0.8$ and consider a range of values for $\gamma$ and $\psi$. We set $R = 1.0033$ for monthly data, corresponding to a 4 percent annual interest rate. Lustig, Stathopoulos and Verdelhan (2017) consider the returns on ten year coupon bonds. A 10-year bond with face value of 1 and coupons of $R - 1 = 0.0033$ has a Macauley duration of 99.3 months or 8.3 years. We set $\delta = 0.0071$, which yields a Macauley duration of 99.3 months. Finally, we calibrate $\Sigma$ as follows. First note that while we focus on interest rate shocks, the volatility of returns is mainly driven by other shocks that we have not explicitly introduced. We use data on real exchange rates and long term bond returns to compute $\Sigma$. Specifically, using the symmetry of the model we can write

$$
\Sigma = \begin{pmatrix}
\sigma_1^2 & \sigma_1^2 - \sigma_{13} & \sigma_{13} \\
\sigma_1^2 - \sigma_{13} & \sigma_1^2 + \sigma_3^2 - 2\sigma_{13} & \sigma_{23} \\
\sigma_{13} & \sigma_{23} & \sigma_3^2
\end{pmatrix}
$$

We therefore compute four moments: the variance of the real exchange rate $\sigma_1^2$, the covariance $\sigma_{13}$ between the real exchange rate and Home real long term bond return, the variance $\sigma_3^2$ of real long term bond returns and the covariance $\sigma_{23}$ between the Home and Foreign real long term bond returns. Details are left to the Data Appendix.
Appendix

A Proof of Lemma 1

It is immediate from the definitions of $\alpha$ and $D$ that they are respectively equal to 0 and $1 + \gamma \sigma^2$ when $\psi = 0$. To show that they both monotonically rise with $\psi$, we take their derivatives:

$$\frac{\partial \alpha}{\partial \psi} = \frac{0.5b}{\theta^2 - 4\psi b} \left( \sqrt{\theta^2 - 4\psi b} - (\theta - 2) \right) \quad (A.1)$$

$$\frac{\partial D}{\partial \psi} = \frac{0.5b}{\theta^2 - 4\psi b} \left( \sqrt{\theta^2 - 4\psi b} + (\theta - 2) \right) \quad (A.2)$$

It is easy to see that $\sqrt{\theta^2 - 4\psi b}$ is larger than both $\theta - 2$ and $2 - \theta$. This is automatic when these are negative. When they are positive, it follows because $\theta^2 - 4\psi b > (\theta - 2)^2$. The latter can be written as $-\psi b > -\theta + 1$, which holds when substituting $\theta = 1 + \psi b + \gamma \sigma^2$.

Next consider the limit of $\psi \to \infty$. We can write

$$\lim_{\psi \to \infty} \alpha = 0.5 \lim_{\psi \to \infty} \frac{1 - \sqrt{1 - \frac{4\psi b}{\theta^2}}}{1/\theta} \quad (A.3)$$

Since both the numerator and denominator approach 0 when $\psi \to \infty$, we can use L’Hopital’s rule:

$$\lim_{\psi \to \infty} \alpha = -0.5 \lim_{\psi \to \infty} \frac{0.5(1/\theta^2)(1 - \frac{4\psi b}{\theta^2})^{-0.5}(-4b + 8\psi b^2/\theta)}{-b/\theta^2} = 1 \quad (A.4)$$

It is immediate that $D \to \infty$ as $\theta \to \infty$ and $D/\theta \to 1$ when $\psi \to \infty$.

B Excess Return Predictability Coefficients

We will now derive the excess return predictability coefficients

$$\beta_k = \frac{\text{cov}(er_{t+k}, r^D_t)}{\text{var}(r^D_t)} \quad (B.1)$$
From \( q_t = \alpha q_{t-1} + \frac{1}{D-\rho} r_t^D \), we have

\[
q_t = \frac{1}{D-\rho} \left( r_t^D + \alpha r_{t-1}^D + \alpha^2 r_{t-2}^D + ... \right) \quad (B.2)
\]

Therefore

\[
er_{t+k} = q_{t+k} - q_{t+k-1} + r_{t+k-1}^D = \frac{1}{D-\rho} r_{t+k}^D + \frac{1}{D-\rho} \left( \alpha - 1 \right) \left( r_{t+k-1}^D + \alpha r_{t+k-2}^D + \alpha^2 r_{t+k-2}^D + ... \right) + r_{t+k-1}^D \quad (B.3)
\]

Then

\[
\text{cov}(q_{t+k} - q_{t+k-1} + r_{t+k-1}^D, r_t^D) = \frac{1}{D-\rho} \rho^k \text{var}(r_t^D) + \rho^{k-1} \text{var}(r_t^D) + \frac{1}{D-\rho} \left( \alpha - 1 \right) \left( \rho^{k-1} + \alpha \rho^{k-2} + ... + \alpha^{k-2} \rho + \frac{\alpha^{k-1}}{1 - \alpha \rho} \right) \quad (B.4)
\]

It follows that

\[
\beta_k = \frac{1}{D-\rho} \rho^k + \rho^{k-1} + \frac{1}{D-\rho} \left( \alpha - 1 \right) \left( \rho^{k-1} + \alpha \rho^{k-2} + ... + \alpha^{k-2} \rho + \frac{\alpha^{k-1}}{1 - \alpha \rho} \right) \quad (B.5)
\]

Consider the last term, but not including the ratio at the end of the large bracketed term. We can rewrite this as

\[
\frac{\alpha - 1}{D-\rho} \alpha^{k-1} \left( \left( \frac{\rho}{\alpha} \right)^{k-1} + \left( \frac{\rho}{\alpha} \right)^{k-2} + ... + \left( \frac{\rho}{\alpha} \right) \right) \quad (B.6)
\]

When \( \alpha \neq \rho \), we can write it as

\[
\frac{\alpha - 1}{D-\rho} \alpha^{k-1} \left( \frac{\rho}{\alpha} \right)^{k-1} - \frac{\alpha - 1}{D-\rho} \alpha^{k-1} \left( \frac{\rho}{\alpha} \right) \quad (B.7)
\]

which can be written as

\[
\frac{\alpha - 1}{D-\rho} \rho \frac{\rho^{k-1} - (\frac{\rho}{\alpha})^k}{1 - (\frac{\rho}{\alpha})} \quad (B.8)
\]

Adding to this the remaining terms of (B.5), we obtain the expression (25) for \( \beta_k \) in the text when \( \alpha \neq \rho \). When \( \alpha = \rho \), (B.6) is equal to

\[
\frac{\alpha - 1}{D-\rho} (k-1) \alpha^{k-1} \quad (B.9)
\]

Adding to this the remaining terms of (B.5), we obtain the expression (25) for \( \beta_k \) when \( \alpha = \rho \).
C Proof of Lemma 2

First consider $\lambda_1$. Since $D - \rho > 0$, the sign is determined by

$$D - \rho \frac{\alpha - 1}{\alpha - \rho} \quad (C.1)$$

$\tilde{\psi}_1$ is defined such that this term is equal to 0. To see this, setting (C.1) equal to zero and substituting the expressions (16) and (17) for $\alpha$ and $D$, we have

$$\frac{\theta + \sqrt{\theta^2 - 4\psi b}}{2} = \frac{\rho - \sqrt{\theta^2 - 4\psi b - 2}}{\theta - \sqrt{\theta^2 - 4\psi b - 2\rho}} \quad (C.2)$$

Cross multiplying delivers

$$\psi b = \rho \theta - \rho \quad (C.3)$$

Substituting $\theta = 1 + \psi b + \gamma \sigma^2 b$ gives $\psi = (\rho/(1 - \rho))\gamma \sigma^2 = \tilde{\psi}_1$.

Now go back to (C.1). It is immediate that this term is positive when $\alpha > \rho$, so that $\lambda_1 > 0$. This happens when $\psi > \tilde{\psi}_2$. So we need to consider $\psi < \tilde{\psi}_2$, so that $\alpha < \rho$. Consider $D$ and $\rho(\alpha - 1)/(\alpha - \rho)$ as functions of $\psi$. It follows from Lemma 1 that both rise monotonically with $\psi$. At $\psi = 0$, so that $\alpha = 0$, $D > \rho(\alpha - 1)/(\alpha - \rho)$. But $\rho(\alpha - 1)/(\alpha - \rho)$ rises to infinity as $\alpha$ approaches $\rho$ from below, which happens when $\psi$ approaches $\tilde{\psi}_2$ from below. Therefore the schedule for $\rho(\alpha - 1)/(\alpha - \rho)$ must cross that for $D$ between $\psi = 0$ and $\psi = \tilde{\psi}_2$. This happens at $\psi = \tilde{\psi}_1$. It follows that $\lambda_1 > 0$ when $\psi < \tilde{\psi}_1$, $\lambda_1 = 0$ when $\psi = \tilde{\psi}_1$ and $\lambda_1 < 0$ when $\tilde{\psi}_1 < \psi < \tilde{\psi}_2$.

Next consider $\lambda_2$. It is immediate from (27) that $\lambda_2 < 0$ when $\alpha > \rho$, which happens when $\psi > \tilde{\psi}_2$. So consider $\psi < \tilde{\psi}_2$, so that $\alpha < \rho$. (27) then implies that $\lambda_2 > 0$ when $1/(1 - \alpha \rho) < \rho/(\rho - \alpha)$. Cross multiplying, this gives $\alpha > \rho^2$. This holds as long as $\alpha > 0$ or $\psi > 0$. When $\psi = 0$, $\alpha = 0$ and $\lambda_2 = 0$.

D Proof of Proposition 1

(21) shows the impulse response to an interest rate shock. First assume $\alpha \neq \rho$. If the interest rate shock starts at time $t = 0$, and we normalize the shock to $D - \rho > 0$ without loss of generality, it implies that in response to this shock

$$q_t - q_{t-1} = \frac{(1 - \rho)\rho^t - (1 - \alpha)\alpha^t}{\alpha - \rho} \quad (D.1)$$
This implies that \( q_1 - q_0 = \alpha + \rho - 1 \). More generally, \( q_t < q_{t-1} \) when
\[
t > \bar{t} = \frac{\ln(1-\rho) - \ln(1-\alpha)}{\ln(\alpha) - \ln(\rho)} \tag{D.2}
\]
while \( q_t > q_{t-1} \) when \( t < \bar{t} \). Below we show that \( \partial \bar{t} / \partial \alpha > 0 \). Since \( \bar{t} = 1 \) when \( \alpha = 1 - \rho \), it follows that \( \bar{t} < 1 \) when \( \alpha + \rho < 1 \). The condition (D.2) is therefore satisfied for all \( t \geq 1 \), so that \( q_t < q_{t-1} \) for all \( t \geq 1 \). This proves the first part of Proposition 1. When \( \alpha + \rho > 1 \), \( \partial \bar{t} / \partial \alpha > 0 \) implies that \( \bar{t} > 1 \). Therefore the real exchange rate continues to appreciate for at least one additional period after the shock \( (t = 1) \), and will start to depreciate once \( t > \bar{t} > 1 \). Finally, when \( \alpha = \rho \), we have \( q_t - q_{t-1} = \rho^{t-1}(\rho - (1 - \rho)t) \) and the same results as those above apply with \( \bar{t} = \rho/(1 - \rho) \). In this case \( \alpha + \rho < 1 \) corresponds to \( \rho < 0.5 \), where \( \bar{t} < 1 \), and \( \alpha + \rho > 1 \) implies \( \rho > 0.5 \), so that \( \bar{t} > 1 \).

It remains to show that \( \partial \bar{t} / \partial \alpha > 0 \) when \( \alpha \neq \rho \). We have
\[
\frac{\partial \bar{t}}{\partial \alpha} = \frac{1}{\alpha(1-\alpha)} \frac{\alpha \ln \alpha + (1-\alpha)\ln(1-\alpha) + \alpha(\ln(1-\rho) - \ln(\rho)) - \ln(1-\rho)}{[\ln(\alpha/\rho)]^2} \tag{D.3}
\]
The sign is determined by the numerator in the large fraction. Note that it is positive for \( \alpha = 0 \) and \( \alpha = 1 \). The derivative of the numerator with respect to \( \alpha \) is \( \ln(\alpha/\rho) - \ln(1-\alpha)/(1-\rho) \), which is positive when \( \alpha > \rho \), zero when \( \alpha = \rho \) and negative when \( \alpha < \rho \). The numerator of the large expression in (D.3) is therefore smallest when \( \alpha = \rho \), where it is zero. It is therefore positive for all \( \alpha \neq \rho \).

**E Proof of Proposition 2**

We have
\[
\beta_1 = \lambda_1 + \lambda_2 = \frac{1}{D - \rho} \left( D - \frac{1-\alpha}{1-\alpha \rho} \right) = \frac{1}{D - \rho} \frac{1}{1-\alpha \rho} (D - \alpha D \rho - 1 + \alpha) \tag{E.1}
\]
Using that \( \alpha D = \psi b \) and \( \alpha + D = \theta \), we have
\[
\beta_1 = \frac{1}{D - \rho} \frac{1}{1-\alpha \rho} (\theta - 1 - \rho \psi b) = \frac{1}{D - \rho} \frac{1}{1-\alpha \rho} (\psi b + \gamma \sigma^2) b > 0 \tag{E.2}
\]
Next consider the second part of Proposition 2. When \( \psi = 0 \), we have \( \alpha = 0 \), \( \theta = 1 + \gamma \sigma^2 b \) and \( D = \theta \). The second part of Proposition 2 then holds when
\[
\frac{1}{D - \rho} \frac{1}{1-\alpha \rho} (\gamma \sigma^2 + (1 - \rho)\psi) > \frac{1}{1 + \gamma \sigma^2 b - \rho} \gamma \sigma^2 \tag{E.3}
\]
This implies

$$(D - \rho)(1 - \alpha \rho) \gamma \sigma^2 < (\gamma \sigma^2 + (1 - \rho)\psi)(1 + \gamma \sigma^2 b - \rho) \quad (E.4)$$

Collecting terms multiplying $\gamma \sigma^2$ and using $D \alpha = \psi b$, we have

$$(D - \psi b \rho + \alpha \rho^2 - 1 - \gamma \sigma^2 b - (1 - \rho)\psi b) \gamma \sigma^2 < (1 - \rho)^2 \psi \quad (E.5)$$

Using $D = \theta - \alpha = 1 + \psi b + \gamma \sigma^2 b - \alpha$, this becomes

$$-\alpha(1 - \rho^2) \gamma \sigma^2 < (1 - \rho)^2 \psi \quad (E.6)$$

which clearly holds.

### F Proof of Proposition 3

The first part of proposition 3 follows immediately from Lemma 2. When $\psi = 0$, we have $\beta_k = \lambda_1 \rho^{k-1}$, which is positive ($\lambda_1 > 0$) and monotonically declines to zero as $k$ rises. When $0 < \psi < \bar{\psi}_1$, Lemma 2 says that both $\lambda_1$ and $\lambda_2$ are positive. Since $0 < \alpha < 1$, it follows that $\beta_k = \lambda_1 \rho^{k-1} + \lambda_2 \alpha^{k-1}$ is positive and monotonically declines to zero with an increase in $k$. Finally, when $\psi = \bar{\psi}_1$, Lemma 2 implies that $\beta_k = \lambda_2 \alpha^{k-1}$, with $\lambda_2 > 0$ and $0 < \alpha < 1$. It again follows that $\beta_k$ is positive and declines monotonically to zero as $k$ rises.

Next consider the second part of Proposition 3, where $\psi > \bar{\psi}_1$. It is immediate from (25) that $\lim_{k \to \infty} \beta_k = 0$. When $\psi \neq \bar{\psi}_2$, so that $\alpha \neq \rho$, we can write

$$\frac{\beta_k}{\alpha^{k-1}} = \lambda_1 \left(\frac{\rho}{\alpha}\right)^{k-1} + \lambda_2 \quad (F.1)$$

$$\frac{\beta_k}{\rho^{k-1}} = \lambda_2 \left(\frac{\alpha}{\rho}\right)^{k-1} + \lambda_1 \quad (F.2)$$

The sign of $\beta_k$ corresponds to the sign of either of the two right hand side expressions. Assume first that $\bar{\psi}_1 < \psi < \bar{\psi}_2$, so that $\alpha < \rho$, $\lambda_1 < 0$ and $\lambda_2 > 0$ (Lemma 2). Then (F.1) implies that $\beta_k > 0$ when $k < \bar{k}_1$ and $\beta_k < 0$ when $k > \bar{k}_1$ with

$$\bar{k}_1 = 1 + \frac{\ln(-\lambda_2/\lambda_1)}{\ln(\rho/\alpha)} \quad (F.3)$$

We know from Proposition 2 that $\beta_1 = \lambda_1 + \lambda_2 > 0$, so that $\lambda_2 > -\lambda_1$, which implies that $\bar{k}_1 > 1$. The $\bar{k}$ in Proposition 3 is the first whole number larger than $\bar{k}_1$. 

30
A similar reasoning applies to the case where $\psi > \bar{\psi}$, so that $\alpha > \rho$, $\lambda_1 > 0$ and $\lambda_2 < 0$ (Lemma 2). Then (F.2) implies that $\beta_k > 0$ when $k < \bar{k}_2$ and $\beta_k < 0$ when $k > \bar{k}_2$ with

$$\bar{k}_2 = 1 + \frac{\ln(-\lambda_1/\lambda_2)}{\ln(\alpha/\rho)} \quad (F.4)$$

From Proposition 2, $\lambda_1 > -\lambda_2$, so that $\bar{k}_2 > 1$. Again the $\bar{k}$ in Proposition 3 is the first whole number larger than $\bar{k}_2$.

Finally consider the special case of $\psi = \bar{\psi}$, so that $\alpha = \rho$. In that case (25) implies that $\beta_k > 0$ when $k < \bar{k}_3$ and $\beta_k < 0$ when $k > \bar{k}_3$ with

$$\bar{k}_3 = 1 + \frac{D - (1/(1 + \rho))}{1 - \rho} > 1 \quad (F.5)$$

Again the $\bar{k}$ in Proposition 3 is the first whole number larger than $\bar{k}_3$.

G Proof of Proposition 4

The Engel condition is $\sum_{k=1}^{\infty} \beta_k < 0$. We will focus here on $\alpha \neq \rho$, which is sufficient as the $\beta_k$ are continuous at $\alpha = \rho$. Then

$$\sum_{k=1}^{\infty} \beta_k = \lambda_1 \frac{1}{1 - \rho} + \lambda_2 \frac{1}{1 - \alpha} = \frac{1}{1 - \rho} - \frac{1}{(D - \rho)(1 - \alpha \rho)} \quad (G.1)$$

The Engel condition can therefore be written as $(D - \rho)(1 - \alpha \rho) < 1 - \rho$. Using that $D\alpha = \psi b$ and $D = \theta - \alpha$, we can also write it as

$$\alpha > \frac{\psi b}{1 + \rho} + \frac{\phi}{1 - \rho^2} \quad (G.2)$$

where $\phi = \gamma \sigma^2 b$. Using $\theta = 1 + \psi b + \phi$ and the definition of $\alpha$, this becomes

$$\sqrt{(1 + \psi b + \phi)^2 - 4\psi b} < 1 - \frac{(1 + \rho^2)\phi}{1 - \rho^2} - \frac{1 - \rho}{1 + \rho} b$$

We can, for convenience, refer to the left and right hand sides of (G.3) as $f(\psi)$ and $g(\psi)$. $f(\psi)$ is a convex function, which is always positive and is symmetric around the axis $\psi = (1 - \phi)/b$, where it reaches a minimum. $g(\psi)$ is a line with a negative slope. Moreover $f(0) > g(0)$. These properties imply that there are only two possibilities. Either $f(\psi)$ remains above $g(\psi)$ for all $\psi$ and therefore the Engel condition is never satisfied, or $f(\psi)$ crosses $g(\psi)$ twice and the Engel
condition is satisfied for an intermediate range of $\psi$ that we will refer to as the interval $(\bar{\psi}_1^E, \bar{\psi}_2^E)$, with the boundaries of the interval equal to the solutions to $f(\psi) = g(\psi)$.

To consider the solutions of $f(\psi) = g(\psi)$, we square both sides. We need to be careful doing so. If $f^2(\psi) = g^2(\psi)$ has two solutions, it is either the case that $f(\psi) = g(\psi)$ for both solutions or $f(\psi) = -g(\psi)$ for both solutions. We know that $f(\psi)$ is convex with an axis of symmetry $\psi = (1 - \phi)/b$. An upward sloping line that crosses $f(\psi)$ twice, will have two solutions with an average value larger than $(1 - \phi)/b$. A negatively sloping line that crosses the symmetric $f(\psi)$ twice will have two solutions that average to less than $(1 - \phi)/b$. Since $g(\psi)$ is a negatively sloping line, the latter must be the case. Otherwise we are finding solutions to $f^2(\psi) = g^2(\psi)$ where $f(\psi) = -g(\psi)$ as $-g(\psi)$ is an upward sloping line. In that case there do not exist any solutions to $f(\psi) = g(\psi)$ and the Engel condition is never satisfied.

We can write $f^2(\psi) = g^2(\psi)$ as

$$A\psi^2 + B\psi + C = 0$$  \hspace{1cm} (G.4)

where

$$A = \rho b^2$$  \hspace{1cm} (G.5)

$$B = b\rho(\phi - 1 - \rho)$$  \hspace{1cm} (G.6)

$$C = \frac{\phi}{(1 - \rho)^2} (1 - \rho^2 - \rho^2\phi)$$  \hspace{1cm} (G.7)

In order for the Engel condition to be satisfied over some intermediate range $(\bar{\psi}_1^E, \bar{\psi}_2^E)$ for $\psi$, two conditions need to hold. First, as discussed above, it must be the case that the average of these solutions is less than $(1 - \phi)/b$, which implies $\phi < 1 - \rho$. Second, it must the case that two solutions to $f^2(\psi) = g^2(\psi)$ exist, which requires $B^2 - 4AC > 0$, which can be written as

$$\rho\phi^2 - 2(2 - \rho)(1 - \rho)\phi + \rho(1 - \rho)^2 > 0$$  \hspace{1cm} (G.8)

This is a quadratic that is positive when $\phi = 0$, then turns negative and then positive again. When $\phi = 1 - \rho$, the quadratic is negative, so that both $\phi < 1 - \rho$ and (G.8) will be satisfied when $\phi$ is between zero and the smaller of the two solutions to (G.8) as an equality. The latter is equal to

$$\bar{\phi} = \frac{1 - \rho}{\rho} \left(1 - \sqrt{1 - \rho}\right)^2$$  \hspace{1cm} (G.9)
To summarize, the Engel condition is satisfied if and only if $\phi < \bar{\phi}$ and $\bar{\psi}_1^E < \psi < \bar{\psi}_2^E$, where $\bar{\psi}_1^E$ and $\bar{\psi}_2^E$ are the solutions to the quadratic (G.4).
References


