Institute for
New Economic Thinking

Working Papers
ineteconomics.org/workingpapers

New Evidence on the Portfolio Balance Approach to Currency Returns

Nevin Cavusoglu[a], Michael D. Goldberg[b,d], and Josh Stillwagon[c,d]

Working Paper No. 89

January 2019

Abstract

This paper re-examines the empirical performance of the portfolio balance approach to currency returns. It considers the implications of two alternative specifications of preferences: one based on expected utility theory and the other on prospect theory. It also uses survey data to estimate models of ex-ante rather than ex-post returns. The empirical analysis relies on the co-integrated VAR framework, which is well suited for testing competing models and dealing with unit roots. Like earlier studies, we find little support for the expected utility theory model. By contrast, the prospect theory model's predictions are largely borne out in the data, including those about sign reversals. We find the strongest support for a hybrid model that incorporates the risk factors of both portfolio balance specifications.

JEL Codes: F31, D81, D84, G10

[a] Department of Economics, James Madison University, Harrisonburg, VA 22807
[b] Department of Economics, University of New Hampshire, Durham, NH 03820
[c] Economics Division, Babson College, Babson Park, MA 02457
[d] Institute for New Economic Thinking (INET) Program on Imperfect Knowledge Economics (IKE)

Keywords: International CAPM, Prospect Theory, Risk Premium, Co-integrated VAR, Survey Expectations

Acknowledgements: We would like to thank Mostafa Beshkar, Bruce Elmslie, Roman Frydman, Soren Johansen, Katarina Juselius, Morten Tabor, and Le Wang for very helpful comments. We also thank the Institute for New Economic Thinking for generous support of this research.
Correspondence: jstillwagon@babson.edu

1 Introduction

The difficulty of conventional risk premium models to account for excess returns in asset markets is well known.[1] Studies find that expected returns are much too volatile to be explained with a plausible degree of risk aversion. Researchers have considered alternative preference specifications to increase the predicted volatility. But, Mark and Wu (1998) show that the failure of the models may lie at a more basic level. They find that the consumption CAPM is grossly inconsistent with the sign reversals in excess returns that occur in currency markets. Lewis (1995) and others find that the older portfolio-balance models are also unable to explain sign reversals.[2]

In this paper, we re-examine the empirical performance of the portfolio balance approach in currency markets. Our investigation differs from other studies in several significant ways. First, we consider two competing models, one based on Dornbusch's (1983) international CAPM (ICAPM) and the other developed by Frydman and Goldberg (2007, 2013). The latter model follows Dornbusch (1983) in its basic setup, but replaces expected utility theory (EUT) with Kahneman and Tversky's (1979) prospect theory (PT). Researchers have found that alternatives to EUT improve the consumption CAPM's empirical performance.[3] Such alternatives may also improve the empirical performance of portfolio balance models.

The two portfolio balance models imply different risk factors. Both models relate the risk premium on foreign currency to the country's bilateral international debt position (IDP). With EUT, the risk premium also depends on the conditional volatility of returns. But, with prospect theory, the premium depends positively on the gap between the exchange rate and market participants' assessments of its benchmark value. The "gap effect" is intuitive: the more over- or undervalued a currency becomes, the riskier it is for market participants who speculate on a further over- or undervaluation. The two models' predictions for sign reversals also differ. Both models relate the risk premium's sign to the sign of IDP. But, with the PT model, the sign of the risk premium also depends on the risk assessments of the bulls

[1] See Fama (2013) and Cochrane (2011).

[2] For review articles on risk premium models in currency markets, see Lewis (1995) and Engel (1996, 2014).

[3] See Backus et al. (2010) and Bansal and Shaliastovich (2013), which incorporate Epstein-Zin (1989) preferences. See also Barberis, Huang, and Santos (2001), which incorporates loss aversion into a consumption CAPM.

(who take long positions in foreign exchange) relative to those of the bears (who take short positions). This additional factor gives the model greater potential to explain sign reversals.

Our empirical analysis estimates models using *ex ante* currency returns, as in Frydman and Goldberg (2007). We measure these returns with monthly survey data on exchange rate expectations. Other risk premium studies estimate models using *ex post* returns, drawing inference under the rational expectations hypothesis (REH). Inference in these studies, therefore, involves joint tests of the models' predictions concerning expected excess returns and REH's prediction that *ex ante* and *ex post* outcomes differ by white noise errors. There is considerable evidence against REH's white-noise-error prediction, suggesting that the negative results of earlier studies may arise in part from a failure of REH.[4] The use of *ex ante* returns enables us to test directly the competing implications of the portfolio balance approach under EUT and prospect theory without the joint hypothesis problem.

A key issue for the empirical analysis is whether excess returns are stationary. Most studies maintain the assumption that they are. But, there is considerable evidence that excess returns are highly persistent and possibly nonstationary.[5] Frydman and Goldberg (2007) address this problem with single-equation error correction models. In this paper, we rely on the $I(1)$ cointegrated vector autoregression (CVAR) framework (Johansen 1996, Juselius 2006), which is better suited for handling persistent variables.[6] We find that expected excess returns and other variables in the information set are best characterized with unit roots.

We also extend Frydman and Goldberg's (2007) analysis by including measures of exchange rate volatility and bilateral IDP in our information set.[7] This enables us to consider all of the PT model's predictions and those of the international CAPM, and a hybrid model that includes both the gap

[4] Compelling evidence comes from survey data studies. In currency markets, see Frankel and Froot (1987), MacDonald and Torrance (1988), Froot and Frankel (1989), Chinn and Frankel (1994), Cavaglia et al. (1994), Bachetta, Mertens and van Wincoop (2000), Stillwagon (2014), and Goldberg at al. (2018). In currency and other markets, see MacDonald (2000) and Bacchetta and van Wincoop (2007).

[5] See Crowder (1994), Evans and Lewis (1994), Frydman and Goldberg (2007), Johansen et al. (2010), Stillwagon (2014, 2018), and Juselius (2014), Jueslius and Assembacher (2017), and Juselius and Stillwagon (2018).

[6] Stillwagon (2018) uses an $I(2)$ CVAR to deal with nonstationarity.

[7] We use a monthly measure of realized volatility as a proxy for the conditional volatility. See section 3.

and ICAPM risk factors. This is the first paper to compare or combine these two specifications of risk. We test the models' competing predictions with alternative over-identifying, long-run restrictions on the VAR.[8]

To preview our results, we find little support for the ICAPM. We reject the model's main prediction — that the expected excess return moves positively with conditional volatility and IDP — in two of the three currency markets examined. By contrast, we find a positive gap effect in all three currency markets, as predicted by the PT model.[9] In two of these markets, IDP also enters the cointegrating relationship as predicted. We also find that the PT model accounts for sign reversals better than the international CAPM. Interestingly, the CVAR results show the strongest evidence for a hybrid model in which the gap, conditional volatility, and IDP drive expected excess returns.

The remainder of the paper is structured as follows. Section 2 reviews the competing predictions of the portfolio balance approach under expected utility theory and prospect theory. Section 3 discusses the information set and data selections, whereas section 4 outlines the CVAR restrictions that are implied by the risk premium models. Section 5 presents the empirical results. We offer concluding remarks in section 6.

2 Competing Portfolio Balance Models

This section sketches more formally the main testable predictions of the ICAPM, PT and hybrid models. We show in section 4 how these predictions can be tested with over-identifying restrictions on the CVAR.

2.1 The International CAPM

The portfolio balance approach under EUT is well known. The model can be expressed as follows:

$$\widehat{rp}_{t+1} = \rho \widehat{v}^{s}_{t+1} IDP_t \tag{1}$$

[8] In modeling currency returns, Della Corte, Riddiough, and Sarno (2016) examine the role of external debt positions, whereas Menkhoff et al. (2012) consider the importance of volatility. Neither study incorporates both variables as implied by the ICAPM.

[9] Frydman and Goldberg (2007) and Stillwagon (2018) also find evidence of a gap effect in the same three currency markets.

where $\widehat{rp}_{t+1}$ denotes the market's risk premium on foreign exchange at time t for $t+1$ (the overhat "$\widehat{\ }$" denotes a forecast), $\widehat{v}^s_{t+1}$ is the market's forecast of the variance of next-period's return, $\rho \geq 0$ is the coefficient of risk aversion, $IDP_t = \frac{B_t^A - A_t^B/s_t}{W_t}$ is the international debt position of country B (the foreign country) vis-a-vis country A (the home country) expressed as a proportion of the total nonmonetary wealth held by both countries, W_t, B_t^A and A_t^B/s_t denote the foreign currency values of B and A bonds that are held entering period t by country A and B wealth holders, respectively, and s_t is the exchange rate (domestic currency price of foreign exchange).[10]

The portfolio balance model in equation (1) implies that the market requires a risk premium on foreign exchange if the foreign country is a net debtor (that is, $IDP_t > 0$). The logic is straightforward. A positive IDP_t implies that the market must hold a net long position in foreign exchange (a short position if $IDP_t < 0$).[11] In equilibrium, market participants on average expect to receive a premium on this open position to compensate them for its riskiness.

The model generates two main testable equilibrium predictions:

1. $\widehat{rp}_{t+1}$ depends positively on the riskiness of holding open positions in foreign exchange, as measured by $\widehat{v}^s_{t+1}$, and the size of the net long position in foreign exchange that must be held (i.e., IDP_t); and

2. $\widehat{rp}_{t+1}$ undergoes a sign reversal when the two countries trade positions as net debtor and net creditor.

Researchers typically generalize equation (1) for the case of multiple foreign countries. REH implies a restriction on the variance-covariance matrix of currency depreciation (Frankel, 1982), which in terms of equation (1) sets $\widehat{v}^s_{t+1}$ equal to the variance of the regression error. It is unclear whether

[10]Equation (1) assumes that domestic and foreign money market equilibrium is determined independently of the spot-rate process and the level of wealth, and that domestic (foreign) money is held only by domestic (foreign) wealth holders. Branson and Henderson (1985) include these assumptions in their "basic asset market specification". We also follow much of the literature and assume that goods prices are nonstochastic.

[11]A positive B_t^A, which includes both inside and outside B bonds, represents the aggregate long position in foreign exchange held by country A wealth holders net of any short positions at time t. Similarly, a positive A_t^B, which also includes both inside and outside bonds, represents the aggregate short position held by country B wealth holders net of any long positions at time t.

Frankel's (1982) and other studies' negative findings are due to a failure of the portfolio balance model or REH. We avoid this joint-hypothesis problem by estimating equation (1) using *ex ante* returns.

2.1.1 The PT Model

The portfolio balance model under prospect theory assumes that market participants are endogenously loss averse: individuals' degree of loss aversion grows with the size of their speculative positions. With endogenous loss aversion, individuals require a risk premium to hold speculative positions in the market in order to compensate them for their extra sensitivity to losses. In the aggregate, the market's risk premium on foreign exchange can be expressed as follows:

$$\widehat{rp}_{t+1} = \frac{1}{2}(\widehat{up}^{\mathrm{L}}_{t+1} - \widehat{up}^{\mathrm{S}}_{t+1}) + \lambda IDP_t \tag{2}$$

where $\lambda > 0$ is a preference parameter and $\widehat{up}^{\mathrm{L}}_{t+1} > 0$ and $\widehat{up}^{\mathrm{S}}_{t+1} > 0$ are aggregations of the minimum expected returns that the bulls and bears, respectively, require in order to hold a speculative position in the market (the superscripts "L" and "S" denote long and short positions, respectively).[12]

Frydman and Goldberg model bulls' and bears' minimum expected returns, which they call "uncertainty premiums," by relating them to the gap between the exchange rate and market participants' assessments of the benchmark value. This assumption builds on an insight from Keynes's (1936) that as an asset becomes more over- or undervalued, it becomes more vulnerable to sharp and sustained counter-movements. The gap, therefore, provides a measure of the riskiness of holding open positions.[13]

Bulls and bears interpret this risk measure in opposite fashion. A greater overvaluation of a currency, for example, exposes bulls (bears) to greater

[12] The bears' premium enters equation (2) negatively because a positive return on a short position entails negative realizations of the excess return. Frydman and Goldberg (2007) assume that the wealth shares of the group of bulls and bears are constant and equal. Hence, the weights of $\frac{1}{2}$ in the equation.

[13] In discussing the decision to hold cash over interest bearing bonds, Keynes (1936) reasoned, "[u]nless reasons are believed to exist why future experience will be very different from past experience, a ...rate of interest [much lower than the benchmark rate], leaves more to fear than to hope, and offers, at the same time, a running yield which is only sufficient to offset a very small measure of fear [of capital loss]" (Keynes, 1936, p.202). See also Tobin (1958).

(lower) potential capital losses. Consequently, bulls raise and bears lower their uncertainty premiums. The portfolio balance equation in (2) shows that both reactions lead to a rise in market's risk premium. This reasoning underpins the following specification for $\widehat{rp}_{t+1}$:

$$\widehat{rp}_{t+1} = \alpha + \gamma gap_t + \lambda IDP_t + \upsilon_t \tag{3}$$

where $gap_t = s_t - s_t^{\mathrm{BM}}$, s_t and s_t^{BM} denote the logarithms of the spot exchange rate and the market's assessment of its benchmark value, respectively, $\gamma > 0$, α is a constant, which when positive (negative) implies that bulls (bears) on average forecast greater potential losses from speculating after accounting for the impact of gap_t, and υ_t is a mean zero error term that represents other factors that may influence bulls' and bears' uncertainty premiums.

Equation (3) shows that a positive risk premium on foreign currency tends to occur when the currency is overvalued ($gap_t > 0$) and/or the foreign country is a net debtor. Both the overvalued currency and debtor status lead bulls to forecast greater potential losses from speculation than the bears and thus to require a greater risk premium for taking speculative positions.

The additional gap term gives the model greater potential to account for sign reversals. Equation (3) shows that the sign of $\widehat{rp}_{t+1}$ depends on the signs and relative magnitudes of $\alpha + \gamma gap_t$ and λIDP_t. The model predicts a sign reversal when the value of $\alpha + \gamma gap_t + \lambda IDP_t$ changes sign. Such reversals can occur when 1) $\alpha + \gamma gap_t$ and λIDP_t are opposite in sign and a switch occurs in the term that takes on the larger absolute magnitude; or 2) the terms are the same sign and both experience a switch in sign. The model also implies that the number of sign reversals should rise as the absolute value of $\alpha + \gamma gap_t + \lambda IDP_t$ falls. A lower absolute value implies a smaller risk premium. As $\widehat{rp}_{t+1}$ gets closer to the zero line, realizations of υ_t and typical changes in the gap_t and IDP_t have a greater chance of triggering a reversal in its sign.

The model implies two sets of testable equilibrium predictions:

1. $\widehat{rp}_{t+1}$ depends positively on both gap_t and IDP_t; and

2. sign reversals tend to occur when $\alpha + \gamma gap_t + \lambda IDP_t$ changes sign; and the number of sign reversals should rise as the size of $\alpha + \gamma gap_t + \lambda IDP_t$ falls.

2.2 The Hybrid Model

The hybrid model is motivated by Barberis, Huang and Santos's (2001) consumption CAPM, which appends the assumption of loss aversion to an otherwise standard specification of preferences based on expected utility theory. The study finds that the hybrid model can resolve the equity premium puzzle.

In similar fashion, we consider a preference specification that concatenates the specifications of the international CAPM and PT model. For statistical reasons, we express the resulting specification for $\widehat{rp}_{t+1}$as follows:[14]

$$\widehat{rp}_{t+1} = \alpha + \gamma gap_t + \rho \widehat{v}^s_{t+1} IDP_t + \upsilon_t \tag{4}$$

where the testable equilibrium predictions derive from the two individual models:

1. $\widehat{rp}_{t+1}$ depends positively on both gap_t and $\widehat{v}^s_{t+1} IDP_t$ and

2. sign reversals tend to occur when $\alpha + \gamma gap_t + \widehat{v}^s_{t+1} IDP_t$ changes sign; and the number of sign reversals should rise as the absolute size of $\alpha + \gamma gap_t + \widehat{v}^s_{t+1} IDP_t$ falls.

3 The Information Set

We estimate the portfolio balance models for three major currency markets, those for the British pound (BP), German mark (DM), and Japanese yen (JY) with respect to the U.S. dollar (USD). Our information set consists of eight variables — $\widehat{\Delta s}_{t+1}, i_t, i^*_t, gap_t, \Delta p_t, \Delta p^*_t, \widehat{v}^s_{t+1}, IDP_t$ — where $\widehat{\Delta s}_{t+1} = \widehat{s}_{t+1|t} - s_t$ denotes the market's time-t forecast of the change in the spot exchange rate from t to $t+1$ (measured as a log difference), i_t, and i^*_t are domestic and foreign nominal interest rates, respectively, Δp_t and Δp^*_t are the domestic and foreign inflation rates (measured again as log differences), respectively, and gap_t, $\widehat{v}^s_{t+1}$, and IDP_t are defined as before. Including $\widehat{\Delta s}_{t+1}$, i_t, and i^*_t as separate variables enables us to test the EUT and PT models' implication that these variables enter the cointegrating system with a $(1, 1, -1)$ restriction, that is, they enter through $\widehat{rp}_{t+1} = \widehat{\Delta s}_{t+1} + i^*_t - i_t$. We

[14]Including IDP_t and $\widehat{v}^s_{t+1} IDP_t$ in the same information set is not desirable, because of multicollinearity.

can also better examine which variables may be adjusting to disequilibrium in the system.

We proxy $\widehat{\Delta s}_{t+1}$ with survey data from Money Market Services International (MMSI). The data are monthly observations on the median exchange rate forecast at the four-week horizon from 1982:11 through 1997:01.[15] We match MMSI's median forecasts with the spot exchange rate that prevailed on the day of each survey (mid-day quotes). MMSI's data compares favorably with other survey data sets in terms of measurement error.[16]

Our measure of the market's benchmark exchange rate relies on purchasing power parity (PPP). To obtain a PPP exchange rate series, we follow Frydman and Goldberg (2007) and use the Big Mac PPP exchange rate as reported in the April 1993 issue of *The Economist*. We update the PPP exchange rate forwards and backwards using domestic and foreign consumer inflation rates. The underlying assumption is that Big Mac inflation follows closely with CPI inflation. The resulting measure of the gap is thus a real exchange rate: $gap_t = s_t - (p_t - p_t^*)$.[17]

Exchange rate data are from Data Resources Inc. (DRIFACS). CPI and interest rate data (10 year bond rate) are from International Financial Statistics.[18]

Empirical researchers typically find that exchange rates, goods prices, and nominal interest rates are $I(1)$. This research relies largely on univariate unit root tests. However, Juselius (2006, 2014) shows that restrictions imposed

[15] MMSI began its surveys in 1982:11 and stopped in 1997:01. Prior to 1985:01, MMSI surveyed participants for their two-week rather than four-week forecast.

[16] MMSI and other surveys asked participants for their forecast of s_{t+1} rather than Δs_{t+1}. Consequently, survey measures of $\widehat{\Delta s}_{t+1}$ at each point in time could imply an expected appreciation or depreciation, depending on the spot exchange rate one uses to obtain an expected change. The measurement error is particularly severe for the FX4casts surveys (also called Forecasts Unlimited Inc. and formerly known as Currency Forecasters' Digest), which give participants a four day window to submit their responses each week. By contrast, MMSI completed its surveys in one day, the Friday of each week. Studies that use FX4casts include Frankel and Chinn (1993), Chinn and Frankel (2002), Bacchetta et al. (2009), and Furnagiev and Stillwagon (2018).

[17] Furnagiev and Stillwagon (2018) proxy s_t^{BM} with a moving average of the exchange rate. Using survey data from FX4casts, they also report a positive gap effect for all four dollar markets examined.

[18] The negative results for the international CAPM are unaffected when short-term interest rates are used. However, the evidence in favor of the PT and hybrid models is weaker. A positive gap effect is found with short-term rates, but IDP_t enters the cointegrating vector significantly with the wrong sign in two of the three markets examined.

on the CVAR provide more powerful tests of unit roots, especially when the signal to noise ratio is low. Using this systems approach, she finds that exchange rates, interest rates, and goods prices are better characterized as near $I(2)$, implying highly persistent first differences.[19] Juselius (2006) shows that a nominal to real transformation may help avoid inference problems in the $I(1)$ CVAR analysis. We thus restrict the exchange rate and goods prices to enter the model through gap_t as one variable. In order to avoid a loss of information (about which variables within the real exchange rate are adjusting), we add the inflation rates, Δp_t and Δp_t^*, to the information set (see Juselius, 2006).

Frankel's (1982) estimation of the ICAPM assumed constant conditional variances. Subsequent studies allowed variances to follow an ARCH or GARCH process or to depend on macroeconomic variables.[20] Our empirical analysis also allows for time-varying conditional variances. But, unlike other studies, we proxy $\hat{\nu}_{t+1}$ with a realized measure of volatility based on daily exchange rate returns.[21] Realized measures are found to be unbiased and efficient estimators of volatility.[22] We also use the realized standard deviation rather than the variance of returns, which we base on the preceding two months of daily returns.[23]

Both portfolio balance models relate a country's risk premium to its bilateral international debt position. Monthly bilateral data on IDP_t are not available. To address this problem, we make use of two different proxies for IDP_t. One proxy uses quarterly bilateral data on current account balances, which are now avaliable from the U.S. Bureau of Economic Analysis (BEA).[24] We set IDP_t for each country vis-a-vis the U.S. (which we take to be the foreign country) to zero in the last quarter of 1988, which is when the

[19] Other studies also find near-$I(2)$ behavior in exchange rate and macro data. See Kongsted and Nielsen (2004), Kongsted (2003, 2005), Johansen et al. (2010), Stillwagon (2018), and Juselius and Stillwagon (2018).

[20] See Engel and Rodriguez (1989), Giovannini and Jorion (1989), and Engel (1994).

[21] Intraday data on returns, needed to produce daily realized measures, is not available over the duration of our survey data sample.

[22] See French et.al. (1987) for monthly measures, and Andersen et.al. (2003) and Barndorff-Nielsen and Shephard (2002) for daily measures.

[23] ARFIMA estimates show that realized variance measures are fractionally integrated. This is also the case with realized standard deviation measures that rely on daily returns over one month, but not over two months. Similar CVAR results are obtained using the alternative one-month measures.

[24] We interpolate the quarterly data using the Chow Lin (1971) procedure in RATS.

U.S.'s overall international financial position changed from net creditor to net debtor status. Positive values of IDP_t, therefore, imply that the U.S. is a bilateral net debtor. Observations on IDP_t for each month going forward (backward) from 1988Q4 are obtained by subtracting (adding) the U.S. bilateral current account balance (i.e., the capital inflow) each month. This current account measure of IDP_t captures capital flows of all types and not just those involving government bonds. However, the measure is imperfect. It misses the stock valuation effects due to exchange rate and asset price fluctuations.

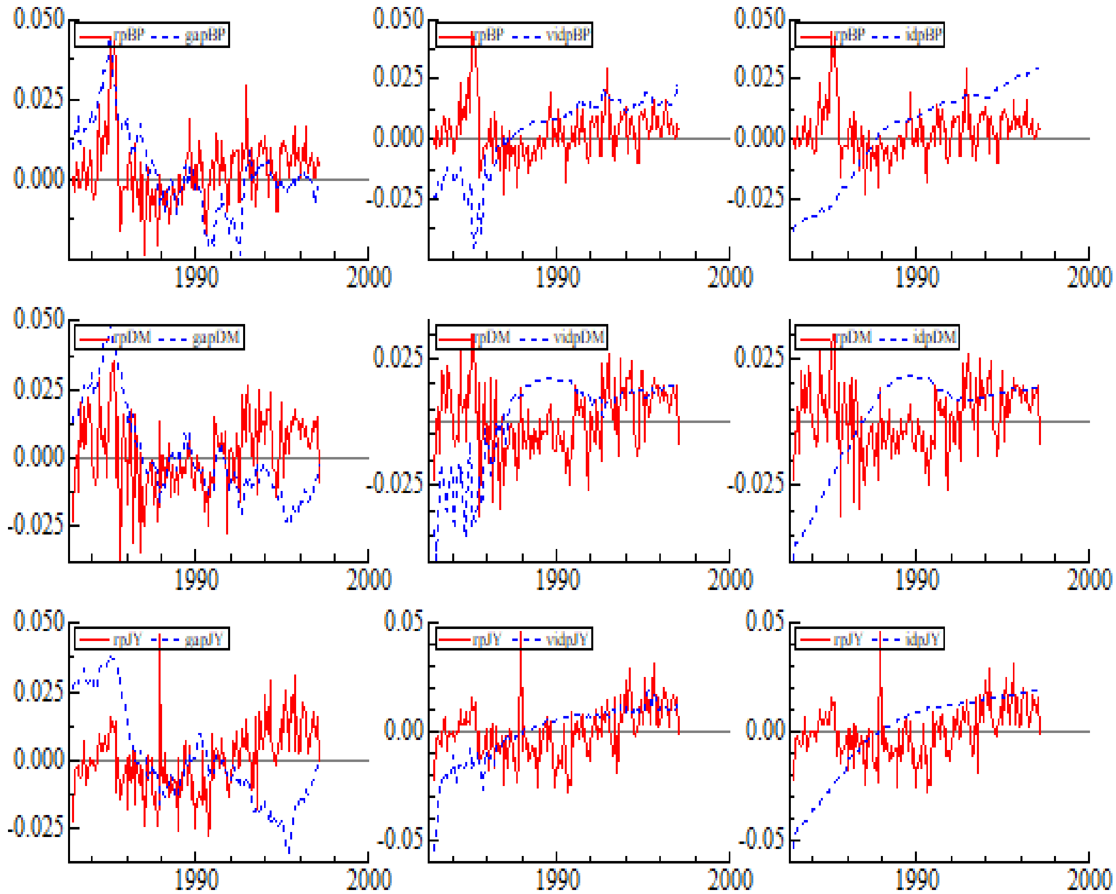
We also proxy the PT model's IDP_t term with a deterministic time trend in the cointegrating space. The U.S.'s international debt position exhibits a rather smooth upward trend in our sample. If the bilateral positions also show such trends (the bilateral current account-measures do), they would be well captured with deterministic trends. This deterministic proxy, which we denote by IDP_t^T, enables us to check whether tests of the model's gap_t prediction may be influenced by IDP_t's highly persistent, near-$I(2)$ behavior.[25]

In Figure 1, we provide simple time plots of our survey-based measure of $\widehat{rp}_{t+1}$ and the hypothesized risk factors. The time plots are illuminating. They show that the risk premium for all three currencies is persistent, undergoing upswings and downswings for extended periods of time. The $\widehat{rp}_{t+1}$ series are also largely positive or negative for protracted periods of time. This behavior is indicative of a highly persistent (possibly unit root) process. It suggests that uncovered interest rate parity may provide a poor characterization of expected excess returns.

Figure 1 also shows that the current account measures of IDP_t, in the third column, display smooth, upward secular trends in all three markets, which is indicative of $I(2)$ behavior. The upward trends appear to have decreased beginning in the 1990s. This suggests that we may need to allow for broken trends when we model IDP with deterministics. The figure also shows in the second column that the ICAPM's risk factor—$\widehat{v}_{t+1}^{s} IDP_t$—shows less signs of $I(2)$ behavior.

[25]The presence of $I(2)$ variables does not influence coefficient estimates in the equilibrium component of the model. However, it can cause biased standard errors, which in general are too small for valid inference. See Juselius (2006) and Franchi and Johansen (2017).

Figure 1: Simple Time Plots



Rows one through three provide series for the BP/USD, DM/USD, and JY/USD markets, respectively. The survey-based measure of the risk premium is depicted in solid red. The other variables are in dotted blue with their means and ranges matched to the premium. The columns provide series for gap_t, $\widehat{v}^s_{t+1} IDP_t$, and IDP_t, respectively, for each market.

The time plots are also suggestive of a positive gap effect: $\widehat{rp}_{t+1}$ tends to co-vary positively with gap_t in all three markets as shown in the first column of Figure 1. For example, the large upswing and downswing in the dollar's value relative to PPP that occurred in the 1980s is associated with a corresponding upswing and downswing in the market's risk premium in all three markets. Connections between $\widehat{rp}_{t+1}$ and IDP_t or $\widehat{v}^s_{t+1}IDP_t$ are much less clear, although there appear to be secular co-movements with both variables, particularly in the 1990s.

4 CVAR Hypotheses

An $I(1)$ CVAR model can be expressed as follows:

$$\Delta x_t = \Gamma \Delta x_{t-1} + \Pi x_{t-1} + \mu_t + \varepsilon_t \tag{5}$$

where x_t denotes the information set, $\Gamma \Delta x_{t-1}$ and Πx_{t-1} are the short-run and equilibrium components of the model, respectively, μ_t is the deterministic component (constant, mean shifts, broken trends, etc.), and ε_t are i.i.d. error terms. Tests of the international CAPM and PT model involve testing over-identifying linear restrictions on Π.

We follow Juselius's (2006) estimation methodology. We first estimate general unrestricted models (GUMs) for each market, which often involves inclusion of dummy variables and mean shifts into the deterministic component in order to obtain well-behaved errors. Ideally, we would want to estimate two GUMs, one for the information set $x_t' = [\widehat{\Delta s}_{t+1}, i_t, i_t^*, gap_t, \Delta p_t, \Delta p_t^*, \widehat{v}^s_{t+1}, IDP_t]$ and the other for $x_t' = [\widehat{\Delta s}_{t+1}, i_t, i_t^*, gap_t, \Delta p_t, \Delta p_t^*, IDP_t^T]$. The second GUM enables us to test the PT model's gap_t predictions without problems associated with IDP_t's high persistence. However, the CVAR is not set up to handle non-linear restrictions, which would be needed to test the ICAPM's prediction that $\widehat{v}^s_{t+1}IDP_t$ enters the cointegrating space positively. We thus estimate three GUMs for each market, one for the ICAPM — $x_t' = [\widehat{\Delta s}_{t+1}, i_t, i_t^*, gap_t, \Delta p_t, \Delta p_t^*, \widehat{v}^s_{t+1}IDP_t]$ and two for the PT model — $x_t' = [\widehat{\Delta s}_{t+1}, i_t, i_t^*, gap_t, \Delta p_t, \Delta p_t^*, IDP_t]$ and $x_t' = [\widehat{\Delta s}_{t+1}, i_t, i_t^*, gap_t, \Delta p_t, \Delta p_t^*, IDP_t^T]$. We refer to the two PT specifications as PT1 and PT2, respectively. The first GUM also enables us to consider a hybrid model in which both gap_t and $\widehat{v}^s_{t+1}IDP_t$ enter the cointegrating space. This hybrid model is motivated by Barberis, Huang and Santos (2001), who find evidence that both EUT and loss aversion are needed to account for returns in equity markets.

The portfolio balance models in equations (1)-(4) represent equilibrium conditions for the foreign exchange market. They imply competing restrictions on Π, which enable us to test which model, if any, can account for fluctuations in the market's risk premium. We also test the implications of two other widely used models of excess returns. One is uncovered interest parity (UIP), which sets $\widehat{rp}_{t+1} = 0$. The other model supposes that the market's risk premium is constant but nonzero, which we call "risk-adjusted UIP" (RAUIP). A necessary (but not sufficient) condition for both UIP and RAUIP is that $\widehat{rp}_{t+1}$ is stationary. In terms of the CVAR, both models imply that the symmetry restrictions underpinning $\widehat{rp}_{t+1}$ give rise to a cointegrating relation, either with a zero or nonzero mean.

Froot and Frankel (1989) and others report survey-based evidence for RAUIP, but do not test the stationarity of $\widehat{rp}_{t+1}$. Direct tests of the stationarity assumption using survey data are few in number.[26] The CVAR provides a powerful approach for such testing. Table 1 summarizes the competing long-run restrictions that are implied by the five models considered in this study.

Table 1: Restrictions on the cointegrating space

	UIP	RAUIP	ICAPM	PT1	PT2	Hybrid
$\widehat{\Delta s}_{t+1}$	$\beta_1 = 1$	$\beta_1 = 1$	$\beta_1 = 1$	$\beta_1 = 1$	$\beta_1 = 1$	$\beta_1 = 1$
i_t^*	$\beta_2 = 1$	$\beta_2 = 1$	$\beta_2 = 1$	$\beta_2 = 1$	$\beta_2 = 1$	$\beta_2 = 1$
i_t	$\beta_3 = -1$	$\beta_3 = -1$	$\beta_3 = -1$	$\beta_3 = -1$	$\beta_3 = -1$	$\beta_3 = -1$
gap_t	$\beta_4 = 0$	$\beta_4 = 0$	$\beta_4 = 0$	$\beta_4 > 0$	$\beta_4 > 0$	$\beta_4 > 0$
IDP_t	-	-	-	$\beta_5 > 0$	-	-
IDP_t^T	-	-	-	-	β_5	-
$v_t IDP_t$	$\beta_5 = 0$	$\beta_5 = 0$	$\beta_5 > 0$	-	-	$\beta_5 > 0$
Δp_t^*	$\beta_6 = 0$	$\beta_6 = 0$	$\beta_6 = 0$	$\beta_6 = 0$	$\beta_6 = 0$	$\beta_6 = 0$
Δp_t	$\beta_7 = 0$	$\beta_7 = 0$	$\beta_7 = 0$	$\beta_7 = 0$	$\beta_7 = 0$	$\beta_7 = 0$
const.	$\beta_8 = 0$	β_8	β_8	β_8	β_8	β_8

[26] Stillwagon (2014, 2015) present CVAR evidence that survey measures of $\widehat{rp}_{t+1}$ are nonstationary.

5 CVAR Results

We first estimate a GUM for each of the three information sets. As with other CVAR studies, we find that the inclusion of dummies and mean shifts is needed in all of the GUMs to obtain valid statistical models. We next test for reduced rank. The trace test shows that Π has reduced rank for all GUMs. We report the results for the hybrid model in the Appendix. The results for the other models are available on request.

We find a rank (which we denote by r) of four for the ICAPM and PT1 specifications and three for the PT2 specification. The number of cointegrating relations for each GUM is equal to its rank.

With $r = 4$ ($r = 3$) we need three (two) restrictions on each cointegrating relation for identification. The international CAPM and the PT and hybrid models imply two symmetry restrictions that constrain the variables $\widehat{\Delta s}_{t+1}$, i_t^*, and i_t to enter the model through $\widehat{rp}_{t+1} = \widehat{\Delta s}_{t+1} + i_t^* - i_t$. They also set the coefficients on the domestic and foreign inflation rate to zero, giving rise to two more restrictions. These four restrictions over-identify one of the cointegrating relations in each GUM. In the remainder of this section, we present the results of testing the competing restrictions in Table 1.

5.1 UIP or RAUIP: Is the Risk Premium Stationary?

Tests of UIP and RAUIP are presented in Table 2 (columns two and three). The figures in the table are the p-values from likelihood ratio tests of the prediction that the corresponding restrictions in Table 1 give rise to a cointegrating (stationary) relation. A Bartlett correction (which raises the p-value) is used in order to address potential small sample bias. The p-value works like an adjusted R-squared that penalizes the inclusion of irrelevant variables. The last column of Table 2 presents results for tests of RAUIP allowing for the mean shifts that we find for the other models in the DM and JY samples. Ignoring these mean shifts could bias the results against stationarity.

The results show rejections of UIP's stationarity prediction at high significance levels for the BP and JY markets. RAUIP is also rejected for these markets (although only marginally for the BP market). By contrast, we cannot reject UIP or RAUIP for the DM market at conventional significance levels. However, we find much stronger evidence of a cointegrating relation in this market when the ICAPM's and PT model's nonstationary risk factors are included in the relation. Even when allowing for a mean shift in the premium, we obtain a markedly higher p-value when including the gap_t (see the next subsection). This finding indicates that $\widehat{rp}_{t+1}$ is also highly persistent in the DM market (and correlated with gap_t).

Taken as a whole, we find that UIP and RAUIP provide poor characterizations of $\widehat{rp}_{t+1}$. The high persistence of $\widehat{rp}_{t+1}$ calls into question the many risk premium studies that maintain the stationarity assumption. As with Mark and Wu (1998), it points to a more fundamental problem for the consumption CAPM, whose second moment risk factor is typically characterized as stationary.

Table 2: Cointegration tests of a stationary risk premium

Exchange Rate	No const.	w/ const.	w/ const. and break
DM	0.171	0.137	0.360
BP	0.024	0.067	
JY	0.000	0.000	0.000

Figures denote p-values for the restrictions of a stationary premium with mean zero, nonzero mean, or a broken mean. The break in mean occurs in 1991:03 for the DM and 1993:01 for the JY sample.

5.2 Portfolio Balance Models' Equilibrium Predictions

By contrast, the risk factors of the portfolio balance models — IDP_t and gap_t — are usually characterized as highly persistent and possibly nonstationary. These models thus have greater potential to account for excess returns.

Table 3 presents the results for all models. The table reports p-values from likelihood ratio tests. Column two provides tests of each model's prediction that the corresponding restrictions in Table 1 give rise to a cointegrating relation. Column three provides tests for the stationarity of the entire system.[27] The table also reports coefficient estimates for the equilibrium relations, with t-statistics in parentheses. The models imply that $(\widehat{s}_{t+1|t} - s_t)$, i_t^*, and i_t enter the equilibrium relation through $\widehat{rp}_{t+1}$. We have thus normalized the

[27] The other cointegrating relationships are just identified with any $r-1$ arbitrary restrictions. These restrictions are chosen by deleting regressors with t-values lower than $|1|$. The results show that the second and third cointegrating relationships can be interpreted as inflation equations, while the fourth (when present) is indicative of how the interest rates tend to co-move across countries. The results may reflect the global business cycle and international spillovers. They are available upon request.

coefficient estimates for these variables to $1, 1$,and -1, respectively, for all models.

5.2.1 ICAPM

Consider first the ICAPM's results. We find that the model's equilibrium relationship cannot be rejected as stationary for all three markets. The model's multiplicative risk factor — $v_t IDP_t$ — is also significant in all three markets. However, the factor enters the equilibrium relation with the wrong sign in two of the three markets (DM and JY).[28] According to these estimates, the market interprets a rise in either the volatility of returns or a country's bilateral indebtedness as signaling less, not greater risk. We find that these odd results are overturned when gap_t is included in the information set.

The results in Table 3 (column 7) also imply nonzero means in the equilibrium relationship. In the DM and JY markets, these means undergo an upward shift after 1991M3 and 1993M1, respectively (column 8).[29] These results are difficult to reconcile with the ICAPM, which implies a zero equilibrium mean when $IDP_t = 0$. They could be an indication that the model is missing one or more important risk factors or that the problem lies with the current account measure of IDP, which omits valuation effects. The results for the PT and hybrid models point towards both interpretations.

5.2.2 PT and Hybrid Models

Consider the PT1 specification, which also uses the current account measure of IDP_t. Table 3 reports higher p-values (columns 2 and 3) for the model than the ICAPM in all three markets, thereby implying stronger evidence of stationary equilibrium relations. This is the case even for the BP sample, where the ICAPM risk factor takes on the correct sign. The higher p-values are an indication that the PT1 model's equilibrium errors are less persistent than those of the ICAPM, suggesting that it provides a better characterization of equilibrium excess returns.

[28] The CVAR results treat all variables as left-hand (potentially endogenous) variables, that is, the cointegrating relation is expressed as $\widehat{rp}_{t+1} + \beta x_t$. Consequently, a negative β estimate for a variable implies a positive relationship with $\widehat{rp}_{t+1}$.

[29] Recall that the equilibrium relationship is expressed as $y + \beta x$. The estimate of the mean after the shift is obtained by adding the figures in columns 7 and 8.

Table 3: Estimates of the cointegrating relationship: $\widehat{rp}_{t+1} + \beta x$

Model	p-value	p-value	gap_t	v_t*IDP_t	IDP_t	constant*	shift
BP/USD	relation	system					
ICAPM	0.282	0.630		−0.361 [−3.211]		−0.002 [−1.822]	
PT1	0.640	0.969	−2.591 [−4.977]		−0.004 [−5.183]	−0.002 [−1.393]	
PT2	0.617	0.617	−1.936 [−2.330]			−0.002 [−2.238]	
Hybrid	0.823	0.823	−2.336 [−2.919]	−1.203 [−6.949]		−0.002 [−1.822]	
DM/USD							
ICAPM	0.237	0.564		1.179 [10.942]		0.009 [5.311]	−0.012 [−5.325]
PT1	0.531	0.531	−2.028 [−2.905]		0.001 [2.613]	0.003 [1.555]	−0.013 [−5.167]
PT2	0.514	0.514	−2.660 [−4.563]			0.001 [0.480]	−0.012 [−4.866]
Hybrid	0.284	0.854	−19.655 [−7.441]	−5.667 [−6.894]		−0.039 [−5.621]	−0.034 [−5.287]
JY/USD							
ICAPM	0.150	0.804		0.274 [8.562]		0.005 [3.539]	−0.019 [−6.579]
PT1	0.781	0.995	−7.046 [−9.526]		−0.002 [−9.236]	0.014 [4.408]	−0.005 [−1.556]
PT2	0.478	0.478	−1.888 [−3.405]			0.001 [0.475]	−0.000 [−2.621]
Hybrid**	0.643	0.686	−1.847 [−4.338]	−0.123 [−8.958]			−0.013 [−4.988]

p-values are for tests of whether the restrictions in Table 1 deliver cointegrating relations (column 2) or a stationary system (column 3). The relation p-values are based on the other just-identified relations. Figures in brackets in columns 4-8 are t-values for the coefficient estimates of the predicted cointegrating relationship within the over-identified system. *Column 7 reports estimates of the mean equilibrium return (the constant) for all models except the PT2 model for the JY market. A mean shift was needed in 1991:03 for the DM market and in 1993:01 for the JY market. Adding the figures in columns 7 and 8 provide the estimate of the mean equilibrium return after the shift. The PT2 model's deterministic trend was significant only for the JY market and only after 1993M3. Column 7 reports the estimate prior to the shift. Adding the figures in columns 7 and 8 provide the coefficient estimate after the shift. **Including an insignificant constant in the

JY/USD hybrid model reduces the relation p-value to 0.124.

In the PT1 model the prediction of a positive gap effect (column 4) is also born out in the data: gap_t enters the equilibrium relation highly significantly with the predicted sign in all three markets.[30] The model's IDP term also enters significantly with the predicted sign in the BP and JY markets. In the DM market, IDP_t enters significantly, but with the wrong sign (though less significantly than in the other two samples). Again, this is an odd result, suggesting that rising indebtedness leads to a falling equilibrium risk premium. This negative result is suggestive that the missing valuation effects in the IDP measure may be more important in the DM market.

Nonzero estimates of a mean equilibrium return in column 7 provide evidence of differential behavior across bulls and bears. The estimate in the BP market is not significant. But, its negative value suggests that bulls on average required a greater equilibrium expected return on holding speculative positions (after accounting for the impact of gap_t) compared with the bears over the sample period. In the DM and JY markets, the coefficient estimates are larger and there is evidence of a shift in 1991M3 and 1993M1, respectively. Estimates of the mean expected return after the break are obtained by adding the figures in columns 7 and 8. The results show that prior to the break, the USD bears required a higher equilibrium expected return, whereas after the break, it was the USD bulls who required a higher expected return.

The PT2 results show that the stationarity and gap_t findings are not the product of IDP_t's inclusion. They continue to show strong evidence of a stationary equilibrium relation and a positive gap effect in all three markets. The deterministic trend measure of IDP is significant only in the JY market and only after 1993M1 (columns 7 and 8 in Table 3). This broken-trend result is suggestive that the growing U.S. indebtedness had a greater positive impact on equilibrium expected returns after the break in 1993. In the other markets, we find a nonzero equilibrium mean instead of a deterministic trend (only after a break in 1991M3 for DM sample). These results are also reported in columns 7 and 8 in Table 3. Again, the negative values suggest that bulls on average required a greater equilibrium expected return than the bears

[30] Recall that all variables are treated as potentially endogenous, so that a negative coefficient estimate for gap_t implies a positive gap effect. Frydman and Goldberg (2007) point out that a positive gap effect runs counter to a house-money effect (Thaler and Johnson, 1990). Barberis, Huang and Santos (2001) use a house money effect to account for stock returns.

(after accounting for the impact of gap_t).

The results for the hybrid model show that the ICAPM's risk factor is after all important, but only after controlling for a gap effect. The results again provide strong support for the PT model's prediction of a positive gap effect in all three markets. But unlike the mixed results for its IDP_t prediction, the ICAPM's $v_t IDP_t$ enters the equilibrium relation significantly with the predicted sign in all three markets. The p-value for the hybrid model's equilibrium relation in (4) is also higher than the p-values for the PT specifications in the BP and JY markets. Taken as a whole, the CVAR results indicate that the hybrid model provides the best characterization of equilibrium excess returns.

5.3 Quick Error Correction

The CVAR's α estimates enable us to examine the short-run dynamics that occur because of deviations from equilibrium relations. They show which variables adjust in the system and whether these adjustments involve error correcting or error increasing behavior. They also shed light on the speed of adjustment.

The portfolio balance models in equations (1)-(4) imply that the exchange rate adjusts to eliminate any disequilibrium in the currency market. The adjustment back to equilibrium works through the exchange rate's impact on $(\widehat{s}_{t+1|t} - s_t)$, gap_t, and IDP_t. The models predict, therefore, that disequilibrium situations will trigger error-correcting movements in these variables. They also predict that this adjustment occurs quickly. In theory, the market should be in equilibrium continuously. In the context of a more general discrete-time empirical model, we would expect that deviations from equilibrium values would be eliminated in one time period barring further shocks.

In general, we would expect that interest rates and goods prices would also be part of the adjustment process. But the partial equilibrium models in equations (1)-(4) assume that these variables are exogenous and determined independently of the exchange rate.

Having rejected the ICAPM's equilibrium predictions in two of the three markets, we focus on the PT and hybrid models. The cointegration results showed the strongest support for the hybrid model. We present the α estimates for this model in Table 4.[31] A significant α estimate for any one of the

[31] The results for the PT1 model show similar short-run dynamics. See the appendix.

seven variables in the system (t-values are in brackets below the coefficient estimates) implies that the corresponding variable adjusts to a disequilibrium. An α estimate for a variable whose sign is opposite to the sign of its β estimate in Table 3 implies error correction, whereas the same sign implies error increasing behavior.

Table 4: Estimates of error-correction to $\left(\hat{s}_{t+1|t} - s_t\right) + i_t^* - i_t - \beta_4 gap_t - \beta_5 v_t IDP_t$

	BP/USD	DM/USD	JY/USD
$\Delta^2 p_t^*$	−0.113 [−1.020]	−0.016 [−2.535]	−0.326 [−1.344]
$\Delta^2 p_t$	0.041 [0.559]	−0.007 [−1.398]	−0.153 [−1.188]
Δi_t^*	0.047 [3.818]	−0.001 [−1.376]	−0.058 [−1.849]
Δi_t	0.020 [1.762]	−0.000 [−0.295]	−0.020 [−1.028]
Δgap_t	0.016 [1.048]	0.002 [1.810]	−0.084 [−3.475]
$\Delta(\widehat{s}_{t+1\|t} - s_t)$	−0.518 [−1.488]	0.022 [0.498]	−1.357 [−1.887]
$\Delta(v_{t+1\|t} * IDP_t)$	0.096 [5.027]	0.031 [5.667]	1.777 [6.503]

t-values are presented in brackets under the error-correction coefficient estimates for the hybrid models predicted cointegrating relationship in the over-identified system.

Consider first the results for the JY market, which are largely consistent with the hybrid model's short-run predictions. We find that the α-estimate for $\Delta(\widehat{s}_{t+1|t} - s_t)$ is significant. Its negative value implies that movements in this variable are error correcting as predicted.[32] The α-estimate for $\Delta(v_{t+1|t} * IDP_t)$ is also significant and its positive value is consistent with error correcting behavior. Adjustment through Δgap_t, however, is significant and consistent with error-increasing behavior. Table 4 shows that the domestic (Japanese) interest rate is exogenous, as the model assumes. However, we also find that the foreign (US) interest rate error-corrects to currency market disequilibrium. The hybrid model thus provides an incomplete characterization of the adjustment process.

[32] Recall that the β estimates for $(s^e_{t+1|t} - s_t)$, i_t^*, and i_t have been normalized to $1, 1,$and -1, respectively.

The CVAR results provide no single measure of the speed of adjustment back to equilibrium. The speed depends on the adjustment of all of the variables. However, the size of the α estimate for the $\Delta(\widehat{s}_{t+1|t} - s_t)$ equation provides some indication of the quickness of adjustment: an α-coefficient equal to -1 would imply that deviations from equilibrium are (ceteris paribus) fully corrected in one time period. Table 4 shows that this α estimate for the JY market is large and greater than unity in absolute terms. A coefficient greater than unity would imply not only quick adjustment, but overshooting in one period. The coefficient, however, is imprecisely estimated. We cannot reject the hypothesis that $\alpha = -1$.

The results for the BP market tell a similar story. The α estimate for $\Delta(\widehat{s}_{t+1|t} - s_t)$ is large in absolute terms and its negative value implies error-correcting behavior. But again, the coefficient is imprecisely estimated; we cannot reject the hypotheses that α equals zero or unity at conventional significance levels. The estimate of -0.518 implies that movements in $(\widehat{s}_{t+1|t} - s_t)$, *ceteris paribus*, eliminate more than half of the disequilibrium in one period. The α estimate for $\Delta(v_{t+1|t} * IDP_t)$ is not nearly as large as we saw for the JY market. But, it is again significant and positive, implying error-correcting behavior. Unlike with the JY market, the α estimate for Δgap_t is small and insignificant.

The results for the DM market are less supportive of the hybrid model's predictions. The α estimate for $\Delta(\widehat{s}_{t+1|t} - s_t)$ is small and insignificant. However, the estimate for $\Delta(v_{t+1|t} * IDP_t)$ is significant and Δgap_t marginally significant, with both error-correcting. The results for the DM market are sensitive to how we account for the influence of IDP. Table A7 in the appendix reports results for the PT1 specification, which show that adjustment through $\Delta(\widehat{s}_{t+1|t} - s_t)$ is large, significant, and error-correcting.

This sensitivity and the drawbacks of the current account measure of IDP suggest that the PT2 specification may provide a better characterization of the dynamics under prospect theory.[33] Table 5 presents the results, which are largely supportive of the PT model. We find that the α estimate for $\Delta(\widehat{s}_{t+1|t} - s_t)$ is significant in all three markets. All three estimates are negative, implying error-correcting behavior. They are also large in absolute terms and more precisely estimated. The point estimates imply quick error correction: movements in $(\widehat{s}_{t+1|t} - s_t)$, *ceteris paribus*, eliminate between 40%

[33] The use of interpolated quarterly data implies that the model is unlikely to capture well any monthly adjustment that may be occurring through IDP.

to 70% of the disequilibrium in one period.

Table 5: Estimates of error-correction to $\hat{s}_{t+1|t} - s_t + i_t^* - i_t - \beta_4 gap_t$

	BP/USD	DM/USD	JY/USD
$\Delta^2 p_t^*$	0.001 [0.056]	0.051 [3.441]	0.072 [3.483]
$\Delta^2 p_t$	−0.011 [−0.468]	−0.118 [−6.392]	−0.275 [−6.521]
Δi_t^*	−0.007 [−2.810]	−0.003 [−1.156]	−0.006 [−1.606]
Δi_t	−0.000 [−0.098]	0.001 [0.400]	0.001 [−0.238]
Δgap_t	−0.003 [−0.908]	0.005 [1.670]	0.004 [0.967]
$\Delta(\widehat{s}_{t+1\|t} - s_t)$	−0.501 [−6.760]	−0.696 [−5.894]	−0.416 [−3.477]

t-values are presented in brackets under the error-correction coefficient estimates for the PT2 model's predicted cointegrating relationship in the overidentified system.

5.4 Sign Reversals

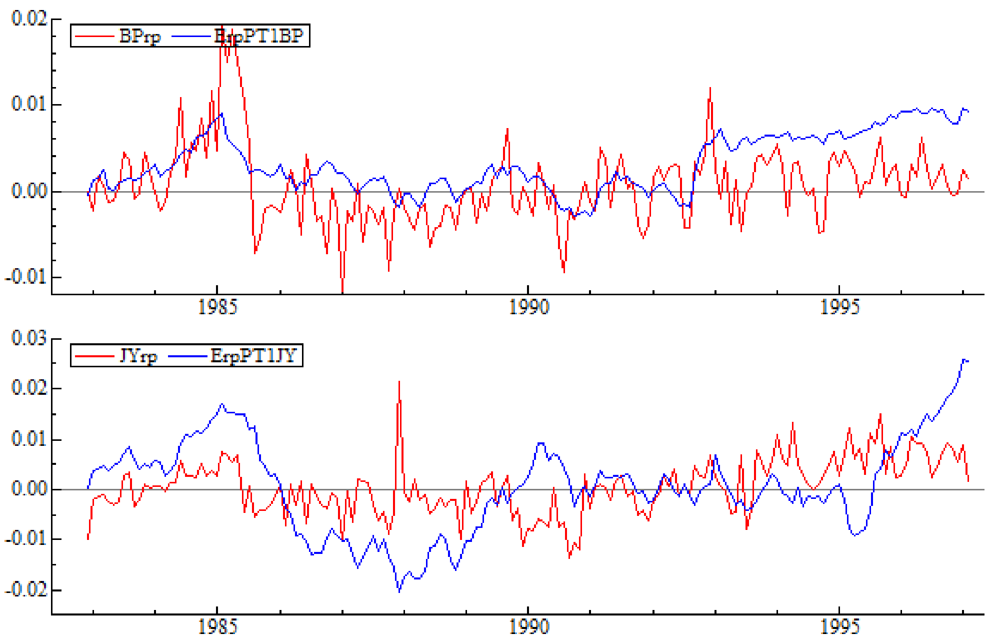
The tendency for the risk premium in currency markets to undergo sign reversals has been a challenge for researchers. Our survey measure of $\widehat{rp}_{t+1}$ undergoes 48, 46, and 54 sign reversals in the BP, DM, and JY markets, respectively, over a sample of 169 monthly observations (see Table 6, column 2). The ICAPM has difficulty accounting for this behavior. IDP_t changes sign only once over the sample in each market. Consequently, even if we ignore rejections of the model's equilibrium predictions in two of the markets, the model predicts only one sign reversal. Mark and Wu (1998) show that the consumption CAPM also has little ability to account for these outcomes.

We saw in section 2 that the gap_t term gives the prospect theory and hybrid models greater ability to account for sign reversals. This ability can be seen in Figures 2 and 3. The figures plot the actual survey risk premium (which we denote by $\widehat{rp}_{t+1}$ and display in dotted red) along with the fitted values of the PT1 and hybrid models' equilibrium relation (which we denote by $\widehat{rp}^{\mathrm{F}}_{t+1}$ and display in solid blue).[34] We found that the PT1 model's equi-

[34] The PT2 specification's ability to account for sign reversals is inferior to the other specifications. This finding provides another indication that IDP_t is an important risk factor in currency markets.

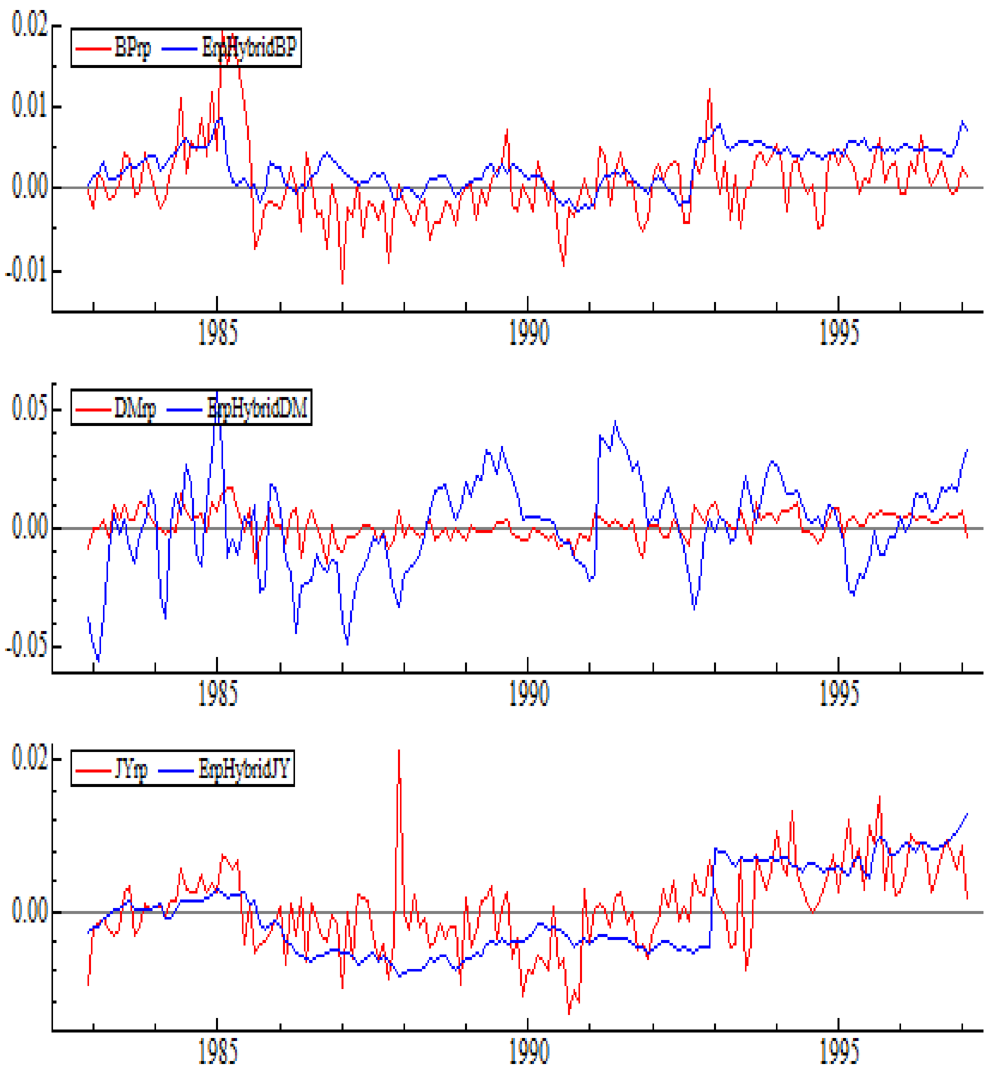
librium prediction concerning IDP_t was rejected in the DM market and so we omit a time plot for this model in Figure 2.

Figure 2: Predicted and Actual Premium for the PT1 Model



The predicted premium is displayed in solid blue while the actual premium series is in dotted red. The upper panel is for the BP/USD sample and the lower panel is for the JY/USD.

Figure 3: Predicted and Actual Premium for the Hybrid Model



The predicted premium is in solid blue and the actual premium series is in dotted red. The upper panel is for the BP/USD sample, the middle is for the DM/USD, and the lower is for the JY/USD.

The figures show that the fitted equilibrium values experience many more sign reversals than predicted by the ICAPM in every market. Table 6 (column 4) reports the exact number of sign reversals exhibited by the estimated models. The results range from a low of 5 sign reversals (or 10% of the total number) for the hybrid model in the JY market to a high of 27 sign reversals (or 59% of the total number) for the hybrid model in the DM market. Table 6 (column 5) also shows that the number of observations for which the sign reversals in both $\widehat{rp}^{\mathrm{F}}_{t+1}$ and $\widehat{rp}_{t+1}$ occur in the same time period is smaller. The number of such contemporaneous sign reversals ranges from a low of 1 for the hybrid model in the JY market (or 2% of the observations) to a high of 8 for the PT1 model in the BP and JY markets (or 17% and 15% of the observations, respectively).

Table 6: Number of actual and predicted sign reversals (SR)

Currency	SR in $\widehat{rp}_{t+1}$	Model	SR in $\widehat{rp}^{\mathrm{F}}_{t+1}$	Contemporaneous SR
BP	48	PT1	13	8
		Hybrid	16	6
DM	46	Hybrid	27	7
JY	54	PT1	22	8
		Hybrid	5	1

As fitted values, we would not expect the $\widehat{rp}^{\mathrm{F}}_{t+1}$ series to exhibit as many sign reversals as the actual $\widehat{rp}_{t+1}$ series or to capture the exact locations. We gauge the explanatory power of the PT1 and hybrid models in two ways.

We first follow Mark and Wu (1998) and examine whether $\widehat{rp}_{t+1}$ tends to take on the same sign as the $\widehat{rp}^{\mathrm{F}}_{t+1}$ series. If the models have explanatory power, the number of observations for which the signs match in the sample should be greater than 50%. We would also expect that this proportion would increase as the size of the $\widehat{rp}^{\mathrm{F}}_{t+1}$ series increased. We consider three thresholds, which we define using the sample standard deviations of the $\widehat{rp}^{\mathrm{F}}_{t+1}$ series.

Table 7 reports the results. The no threshold results show that the models have some ability to account for the sign of the market's risk premium in every market. The results range from a low of 57% for the hybrid model in the DM market to a high of 74% for the hybrid model in the JY market. Even the low figure is significant with 170 observations (the p-value is 0.031). In every market, the tendency for the signs to match increases as the size of the fitted values increases. For the largest $\widehat{rp}^{\mathrm{F}}_{t+1}$ observations, we find figures greater than 80% for both models in the BP and DM markets and greater than 70% for the hybrid model in the DM market.

Table 7: Actual and Fitted Premium of the Same Sign

Currency	Threshold	PT1 model	Hybrid model
BP	none	68	69
	0.5 stdev	82	75
	1 stdev	92	88
	1.5 stdev	100	96
DM	none		57
	0.5 stdev		60
	1 stdev		69
	1.5 stdev		74
JY	none	62	74
	0.5 stdev	70	73
	1 stdev	83	80
	1.5 stdev	80	85
Pooled	none	65	67
	0.5 stdev	76	70
	1 stdev	88	80
	1.5 stdev	90	85

Figures in columns 3 and 4 are the percentage of observations for which $\widehat{rp}_{t+1}$ and $\widehat{rp}^{\mathrm{F}}_{t+1}$ are the same sign. Thresholds are defined using the sample standard deviations of the $\widehat{rp}^{\mathrm{F}}_{t+1}$ series. The pooled sample for the hybrid model combines the observations for all three currency markets, whereas for the PT1 model, it combines the observations for the BP and JY markets only.

We also gauge the models' explanatory power by testing whether the frequency of sign reversals falls as the models' fitted equilibrium value rises and thus predicts a larger market risk premium (see section 2). Tables 8 and 9 report results for the PT1 model and the hybrid model, respectively. We use the same three thresholds as before. The third column in the tables provides the number of observations above or below a threshold. The column labeled "FSR" gives two frequencies for each threshold. The top figure is the number of sign reversals that occur in $\widehat{rp}_{t+1}$ when the absolute value of $\widehat{rp}^{\mathrm{F}}_{t+1}$ is above the threshold, expressed as a proportion of the number of observations above the threshold.[35] The bottom figure is the frequency of sign reversals below the threshold. The "SRR" column shows the ratio of these two frequencies.

Table 8: Frequency of sign reversals: The PT1 model

Currency	Threshold	N	FSR	SRR	Z-value	p-value
BP	0.5 stdev above	85	0.13			
	0.5 stdev below	85	0.44	0.30	4.71	0.000
	1 stdev above	62	0.13			
	1 stdev below	108	0.37	0.35	3.83	0.000
	1.5 stdev above	26	0.00			
	1.5 stdev below	144	0.33	0.00	8.49	0.000
JY	0.5 stdev above	93	0.27			
	0.5 stdev below	77	0.38	0.71	1.50	0.067
	1 stdev above	59	0.25			
	1 stdev below	111	0.35	0.72	1.34	0.090
	1.5 stdev above	25	0.24			
	1.5 stdev below	145	0.33	0.73	0.97	0.166
Pooled	0.5 stdev above	178	0.20			
	0.5 stdev below	162	0.41	0.50	4.19	0.000
	1 stdev above	121	0.19			
	1 stdev below	219	0.36	0.53	3.54	0.000
	1.5 stdev above	51	0.12			
	1.5 stdev below	289	0.33	0.35	4.05	0.000

Thresholds are defined using the sample standard deviations of the $\widehat{rp}^{F}_{t+1}$ series. N in column 3 denotes the number of observations. FSR in column 4 denotes the frequency of sign reversals as a proportion of the corresponding number of observations. SRR in column 5 denotes a ratio of the proportions of sign reversals in $\widehat{rp}_{t+1}$ that are associated with a gap above and below the specified threshold. Column 6 provides a Z statistic for the difference of SRR from 1, and column 7 provides the corresponding p-value. The pooled sample combines the BP and JY markets.

[35] A sign reversal occurs between two adjacent time periods, t and $t+1$. We count a sign reversal as occurring above the threshold if both the t and $t+1$ observations on $\widehat{rp}^{F}_{t+1}$ are above the threshold. The results are slightly less strong if we require just the t observation or just the $t+1$ observation to be above the threshold.

Consider first the BP results. The PT1 and hybrid models' SRR for the smallest threshold are, respectively, .30 and .41, implying that the frequency of sign reversals above the threshold is less than half the frequency below the threshold for both models. A Z-statistic implies that the difference in frequencies is highly significant. The results show that $\widehat{rp}_{t+1}$ undergoes far fewer sign reversals when the $\widehat{rp}^{\mathrm{F}}_{t+1}$ predicts a larger risk premium.

We also find that the models' SRR falls as the size of the threshold rises. At the largest threshold, we find no sign reversals above for the PT1 model. For the hybrid model, the frequency of sign reversals above is nearly 90% lower than below the threshold. A Z-statistic shows that the differences in frequencies are highly significant.

The hybrid model's results in the other markets tell a similar story to the BP results. We find fewer sign reversals above the threshold in both the DM and JY markets at all thresholds. The significance levels are weaker. But, the model's SRR falls at higher thresholds in both markets. At the largest threshold, the frequency of sign reversals above is roughly 80% lower than below the threshold.

The weakest results are found for the PT1 model in the JY market. The model's SRR shows no tendency to fall at higher thresholds. Nonetheless, the ratio is below 1 at all thresholds, but with much weaker significance levels. At the largest threshold, the difference in frequencies is not significant.

The lack of significance in some cases may be due to a small number of observations above or below the threshold. Pooling the observations for each model across all markets avoids this problem. The pooled results tell a consistent story: there is a strong tendency for fewer sign reversals when the models predict a larger risk premium. At the largest threshold, the frequency of sign reversals above the threshold is about a third of the frequency below the threshold for the PT1 model, and a sixth for the hybrid model.

Taken as whole, the results in Tables 7-9 indicate that the PT1 and hybrid models have considerable explanatory power in accounting for sign reversals in currency risk premiums.

Table 9: Frequency of sign reversals: The hybrid model

Currency	Threshold	N	FSR	SRR	Z-value	p-value
BP	0.5 stdev above	101	0.18			
	0.5 stdev below	69	0.43	0.41	3.62	0.000
	1 stdev above	69	0.14			
	1 stdev below	101	0.38	0.39	3.60	0.000
	1.5 stdev above	24	0.04			
	1.5 stdev below	146	0.32	0.13	4.99	0.000
DM	0.5 stdev above	109	0.23			
	0.5 stdev below	61	0.34	0.67	1.58	0.057
	1 stdev above	51	0.18			
	1 stdev below	119	0.31	0.57	1.97	0.024
	1.5 stdev above	23	0.04			
	1.5 stdev below	147	0.31	0.14	4.60	0.000
JY	0.5 stdev above	128	0.31			
	0.5 stdev below	42	0.33	0.94	0.25	0.401
	1 stdev above	74	0.27			
	1 stdev below	96	0.35	0.76	1.18	0.119
	1.5 stdev above	13	0.08			
	1.5 stdev below	157	0.34	0.23	3.14	0.001
Pooled	0.5 stdev above	338	0.25			
	0.5 stdev below	172	0.38	0.65	3.02	0.001
	1 stdev above	194	0.20			
	1 stdev below	316	0.34	0.58	3.66	0.000
	1.5 stdev above	60	0.05			
	1.5 stdev below	450	0.32	0.16	7.62	0.000

Thresholds are defined using the sample standard deviations of the $\widehat{rp}^{F}_{t+1}$ series. N in column 3 denotes the number of observations. FSR in column 4 denotes the frequency of sign reversals as a proportion of the corresponding number of observations. SRR in column 5 denotes a ratio of the proportions of sign reversals in $\widehat{rp}_{t+1}$ that are associated with a gap above and below the specified threshold. Column 6 provides a Z statistic for the difference of SRR from 1, and column 7 provides the corresponding p-value. The pooled sample combines all three currencies.

6 Conclusion

The paper used survey data and the CVAR framework to test directly the competing predictions of several portfolio balance models. The CVAR results rejected UIP's prediction that expected excess returns are stationary in two of the three markets considered. This finding undercuts behavioral models that assume UIP.[36] The results also provided little support to the portfolio balance approach under EUT. The model's risk factor was found to enter the cointegration relation with the wrong sign in two of the three markets.

By contrast, we found support for the PT and hybrid models' equilibrium and short-run predictions. The prediction that market participants use the gap from PPP in assessing the riskiness of speculative positions garnered the strongest evidence: a positive and significant gap effect was found in all three markets for all specifications considered. The CVAR results showed that the market's risk premium also depends on the ICAPM's $v_{t+1|t} * IDP$ risk factor, but only after controlling for the influence of gap_t. We also found that the PT and hybrid models, unlike the ICAPM and consumption CAPM, had explanatory power in accounting for sign reversals in the market's risk premium.

The paper's findings help uncover the source of standard models' empirical difficulties in accounting for *ex post* excess returns: the problem is due in part to the reliance on standard EUT preference specifications. The findings do not address directly the forward rate puzzle, but they are suggestive that the PT and hybrid models would be part of a resolution.

[36] For example, see Mark and Wu (1998) and Gourinchas and Tornell (2004).

References

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Paul Labys (2003). "Modeling and Forecasting Realized Volatility." *Econometrica* 71(2), 579-625.

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

Bacchetta, Philippe and Eric van Wincoop (2007). "Random Walk Expectations and the Forward Discount Puzzle," *American Economic Review*, 97(2), 346-350.

Backus, David K., Federico Gavazzoni, Christopher Telmer, and Stanley E. Zin (2010). "Monetary Policy and the Uncovered Interest Parity Puzzle.". NBER working paper no. w16218 National Bureau of Economic Research.

Bansal, Ravi, and Ivan Shaliastovich (2013)."A Long-run Risks Explanation of Predictability Puzzles in Bond and Currency Markets." *Review of Financial Studies* 26(1), 1-33.

Barberis, Nicholas, Ming Huang and Tano Santos (2001) "Prospect Theory and Asset Prices." *Quarterly Journal of Economics*, 116(1), 1-53.

Barndorff-Nielsen, Ole E. and Neil Shephard (2002). "Estimating Quadratic Variation Using Realized Variance." *Journal of Applied Econometrics* 17(5), 457-477.

Branson, William H., and Dale W. Henderson. "The Specification and Influence of Asset Markets." *Handbook of International Economics* 2 (1985): 749-805.

Cavaglia, Stephano, Willem F.C. VerSchoor, and Christian C.P. Wolff (1993). "Further Evidence on Exchange Rate Expectations." *Journal of International Money and Finance*12 (1), 78-98.

Chinn, Menzie and Jeffrey Frankel (1994). "Patterns in Exchange Rate Forecasts for Twenty-five Currencies." *Journal of Money, Credit and Banking* 26, 757-770.

Chinn, Menzie and Jeffrey Frankel (2002). "Patterns in Exchange Rate Forecasts for Twenty-Five Currencies," *New Developments in Exchange Rate Economics*, (2), 428-439.

Chow, Gregory C. and An-loh Lin (1971). "Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series." *The Review of Economics and Statistics* 53(4), 372-375.

Cochrane, John (2011). "AFA Presidential Address: Discount Rates." *The Journal of Finance* 66(4), 1047-1108.

Crowder, William J. (1994). "Foreign Exchange Market Efficiency and Common Stochastic Trends," *Journal of International Money and Finance*, 13(5), 551-564.

Della Corte, Pasquale, Steven J. Riddiough, and Lucio Sarno (2016). "Currency Premia and Global Imbalances." *The Review of Financial Studies* 29(8), 2161-2193.

Dornbusch (1983). "Exchange Risk and the Macroeconomics of Exchange Rate Determination." In *The Internationalization of Financial Markets and National Economic Policy*, R Hawkins, R Levich and C Wihlborg, Greenwich CT: JAI Press.

Engel, Charles M. (1994). "Tests of CAPM on an International Portfolio of Bonds and Stocks," The internationalization of equity markets, 149-72, National Bureau of Economic Research Project Report series. Chicago and London: University of Chicago Press.

Engel, Charles A. (1996). "The Forward Discount Anomaly and the Risk Premium: A Survey of Recent Evidence." *Journal of Empirical Finance* 3, 123-191.

Engel, Charles A. (2014). "Exchange Rates and Interest Parity" in Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff (eds.), *Handbook of International Economics* Volume 4, Chapter 8.

Engel, Charles and Anthony P. Rodrigues (1989). "Tests of International CAPM with Time-Varying Covariances," *Journal of Applied Econometrics*, 4(2), 119-138.

Epstein, Larry G. and Stanley E. Zin (1989). "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework.". *Econometrica* 57(4), 937-969.

Evans, Martin D. D. and Karen K. Lewis (1994). :Do Stationary Risk Premia Explain It All? Evidence from the Term Structure," *Journal of Monetary Economics*, 33(2), 285-318.

Fama, Eugene (2013). "Two Pillars of Finance", Nobel Prize Lecture, December 8th, 2013.

Franchi, Massimo and Soren Johansen (2017). "Improved Inference on Cointegrating Vectors in the Presence of a Near Unit Root Using Adjusted Quantiles," *Econometrics*, MDPI, Open Access Journal 5(2), 1-20.

Frankel, Jeffrey A. (1982). "A Test of Perfect Substitutability in the Foreign Exchange Market," *Southern Economic Journal*, 49(2), 406-416.

Frankel, Jeffrey A. and Menzie D. Chinn (1993). "Exchange Rate Expectations and the Risk Premium: Tests for a Cross Section of 17 Currencies," *Review of International Economics*, 1(2), 136-144.

Frankel, J. and K. Froot (1987). "Using Survey Data to Test Standard Propositions Regarding Exchange Rate Expectations.". *American Economic Review* 77(1), 133-153.

French, Kenneth R. G. William Schwert, and Robert F. Stambaught (1987). "Expected Stock Returns and Volatility." *Journal of Financial Economics* 19(1), 3-29.

Froot, K. and J. Frankel (1989). "Forward Discount Bias: Is it an Exchange Risk Premium?." *Quarterly Journal of Economics* 104(1), 139-161.

Frydman, R. and M. Goldberg (2007). *Imperfect Knowledge Economics: Exchange rates and Risk*, Princeton, NJ: Princeton University Press.

Frydman, R. and M. Goldberg (2013). "Opening Models of Asset Prices and Risk to Non-routine Change." in Roman Frydman and Edmund S. Phelps (eds.) *Rethinking Expectations: The Way Forward for Macroeconomics*, 207-250.

Furnagiev, Steve and Josh R. Stillwagon (2018). "Currency Risk Premia: Perceptions of Downside Risk and Deviations from Benchmark Values." Working Paper.

Giovannini, Alberto and Philippe Jorion (1989). "The Time Variation of Risk and Return in the Foreign Exchange and Stock Markets," *Journal of Finance*, 44(2), 307-325.

Goldberg, Michael D., Olesia Kozlova, and Peter Sullivan (2018). "Are Currency Returns Really Predictable? Novel Structural Change and the Forward Rate Anomaly.". Working Paper.

Gourinchas, Pierre-Olivier and Aaron Tornell (2004), "Exchange Rate Puzzles and Distorted Beliefs," *Journal of International Economics*, 64, 303-33.

Johansen, Soren, Katarina Juselius, Roman Frydman and Michael Goldberg (2010). "Testing Hypotheses in an I(2) Model with Piecewise Linear Trends: An Analysis of the Persistent Long Swings in the Dmk/$ Rate," *Journal of Econometrics*, 158(1), 117-129.

Juselius, Katarina (2006). *The Cointegrated VAR Model: Methodology and Applications* Oxford University Press.

Juselius, Katarina (2014). "Testing for Near I(2) Trends When the Signal-to-Noise Ratio Is Small," *Economics: The Open-Access, Open-Assessment E-Journal*, 2014, 8, 1-30.

Juselius, Katarina and Katrin Assenmacher (2017). "Real Exchange Rate Persistence and the Excess Return Puzzle: The Case of Switzerland versus the US," *Journal of Applied Econometrics* 32(6), 1145-1155.

Juselius, Katarina and Ronald MacDonald (2004). "International Parity Relationships between Germany and the United States: A Joint Modelling Approach." Finance Research Unit, Institute of Economics, University of Copenhagen.

Juselius, Katarina and Josh R. Stillwagon (2018). "Are Outcomes Driving Expectations or the Other Way Around? An I(2) CVAR Analysis of Interest Rate Expectations in the Dollar/Pound Market." *Journal of International Money and Finance*, 83, 93-105.

Kahneman, Daniel and Tversky, Amos (1979). "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47, 263-291.

Keynes, John M. (1936). *The General Theory of Employment, Interest and Money*, Harcourt, Brace and World: New York.

Kongsted, Hans Christian (2003). "An I(2) Cointegration Analysis of Small-Country Import Price Determination," *Econometrics Journal*, 6(1), 53-71.

Kongsted, Hans Christian (2005). "Testing the Nominal-to-Real Transformation," *Journal of Econometrics*, 124(2), 205-225.

Kongsted, Hans Christian and Heino Bohn Nielsen (2004). "Analysing I(2) Systems by Transformed Vector Autoregressions," *Oxford Bulletin of Economics and Statistics*, 66(3), 379-397.

Lewis, Karen K. (1995). "Puzzles in International Financial Markets," in G. Grossman and K. Rogoff (eds.), *Handbook of International Economics*, Vol. III, Amsterdam: North-Holland,1913-1917.

MacDonald, R. (2000). "Expectations Formation and Risk in Three Financial Markets: Surveying What the Surveys Say.". *Journal of Economic Surveys* 14(1), 69-100.

MacDonald, R. and T.S. Torrance (1988). "On Risk, Rationality and Excessive Speculation in the Deutschmark-U.S. dollar Exchange Market: Some Evidence Using Survey Data." *Oxford Bulletin of Economics and Statistics* 50(2), 107-123.

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

Menkhoff, Lukas, Lucio Sarno, Maik Schmeling, and Andreas Schrimpf (2012). "Carry Trades and Global Foreign Exchange Volatility." *The Journal* *of Finance* 67(2), 681-718.

Rogoff, Kenneth (1996). "The Purchasing Power Parity Puzzle," *Journal of Economic Literature* 34(2), 647-668.

Stillwagon, Josh R. (2014). "Reexamining What Survey Data Says about Currency Risk and Irrationality". *Economics Bulletin* 34(3), 1631-1634.

Stillwagon, Josh R. (2015). "Can the Consumption Capital Asset Pricing Model Account for Traders' Expected Currency Returns?." *Review of International Economics* 23(5), 1044-1069.

Stillwagon, Josh R. (2018). "Are Risk Premia Related to Real Exchange Rate Swings? Evidence from I(2) CVARs with Survey Expectations." *Macroeconomic Dynamics* 22, 255-278.

Thaler, Richard. H. and Eric Johnson (1990), "Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes in Risky Choice," *Management Science*, 36, 643-660.

Tobin, James (1958). "Liquidity Preference as Behavior Towards Risk", *The Review of Economic Studies* 67, 1958.

7 Appendix

7.1 Modeling Specification and Identification

Inference from a CVAR requires a well-specified unrestricted model. The model's results are sensitive to problems of skewness (an absolute value greater than .4 as a rule of thumb) and auto-correlation. By contrast, inference is robust to heteroskedasticity and excess kurtosis (Juselius, 2006). Juselius's methodology is to include impulse dummies, transitory dummies, and/or mean shifts in the model's deterministic component, as needed, to produce well behaved residuals.

As is generally the case with CVAR estimation, we find that the multivariate normality assumption is not satisfied in any of the markets absent intervention dummy variables to control for the largest outliers or shifts in mean. We do not require a mean shift in the BP/USD sample to address serial correlation, but include dummy variables in 1983:07, 1984:09, 1984:12, 1985:02, 1985:03, 1985:04, 1985:06, 1985:07, 1985:09, 1985:10, 1985:11, 1988:01, 1990:01, 1990:03, 1991:04, and 1992:09. The latter is associated with the UK's abrogation of the European Monetary System fixed exchange rate. Many of the other outliers are associated with the peak of the large USD upswing in

1983-85.

For the German mark sample, a mean shift was needed in 1991:03. This volatile period involved German reunification. Juselius and MacDonald (2004), which uses a similar DM/USD information set, but no survey data, also needed dummies and mean shifts to deal with this period. The empirical application herein largely follows their approach. We also need dummy variables in 1983:08, 1984:09, 1984:12, 1985:11, 1986:02, 1987:11, 1989:01, 1990:01, 1990:02, 1990:08, 1991:07, 1991:10, and 1993:01

For the JY/USD sample, we include a mean shift in 1993:01. We also need dummy variables in 1984:03, 1984:08, 1984:12, 1985:09, 1985:11, 1986:02, 1986:03, 1987:12, and 1989:04.

Tables A1 and A2 show that, aside from some second order autocorrelation and skewness in the DM sample's $v * IDP_t$ variable, the GUMs are well-specified in terms of eliminating autocorrelation and skewness.

Table A1: Test for Autocorrelation

	BP/USD	DM/USD	JY/USD
LM(1)	0.058	0.071	0.056
LM(2)	0.532	0.098	0.010

p-value of test for autocorrelation at each lag

Table A2: Skewness by Variable

	BP/USD	DM/USD	JY/USD
$\Delta^2 p_t^*$	0.198	-0.157	0.222
$\Delta^2 p_t$	0.008	0.048	-0.118
Δi_t^*	-0.073	0.315	0.316
Δi_t	-0.231	0.166	0.293
Δgap_t	0.060	0.275	-0.230
$\Delta(\widehat{s}_{t+1\|t} - s_t)$	-0.399	-0.191	-0.098
$v * IDP_t$	0.063	-0.585	-0.130

We use the trace test to determine the rank of the cointegrating space. The test starts from the most restricted model (lowest rank), incrementing r by one until the first clear failure to reject is found. The test results are provided in Table A3. A rank of three is rejected for the BP/USD, and is only a borderline failure to reject for the other two samples. The trace test is unable to reject a rank of four by considerable margins for all three samples.

Table A3: Rank Test Statistics

	BP/USD	DM/USD	JY/USD
r=0	0.000	0.000	0.000
r=1	0.000	0.000	0.000
r=2	0.000	0.000	0.000
r=3	0.010	0.070	0.086
r=4	0.438	0.257	0.992
r=5	0.943	0.419	0.986
r=6	0.972	0.840	1.000

p-values after Bartlett corrections

The roots of the companion matrix provide additional criteria for selecting the appropriate rank of each VAR. Large roots that are close to unity are indicative of common stochastic trends or unit roots in the system, that is, more large roots suggest a lower rank. The number of common stochastic trends is equal to $p - r$, where p is the number of variables and r is the rank. Consequently, the number of large roots, in combination with the number of variables in the information set, can be used to infer the rank. For example, with seven variables in our system, three large roots would imply a rank of four. Tables A4-A6 present the largest roots in each sample for different choices of the imposed rank. We can see that there are four large roots in each sample using Juselius's (2006) "rule of thumb" definition of a root of approximately 0.90 or larger. These large roots are statistically indistinguishable from one in small samples. We can also see that an imposed rank of four eliminates all remaining large roots. Given the trace tests and roots of the companion matrices, a rank of four seems most appropriate in each sample for the hybrid models.

Table A4: BP/USD Largest Roots of the Companion Matrix

r=4	1.000	1.000	1.000	0.830
r=5	1.000	1.000	0.976	0.794
r=6	1.000	0.991	0.957	0.792
Unrestricted VAR	1.000	0.972	0.958	0.814

Table A5: DM/USD Largest Roots of the Companion Matrix

r=4	1.000	1.000	1.000	0.695
r=5	1.000	1.000	0.905	0.696
r=6	1.000	0.915	0.915	0.704
Unrestricted VAR	0.987	0.916	0.916	0.704

Table A6: JY/USD Largest Roots of the Companion Matrix

r=4	1.000	1.000	1.000	0.752
r=5	1.000	1.000	0.946	0.755
r=6	1.000	0.935	0.924	0.758
Unrestricted VAR	0.981	0.965	0.905	0.756

Table A7: Estimates of error-correction to $\left(\hat{s}_{t+1|t} - s_t\right) + i_t^* - i_t - \beta_4 gap_t - \beta_5 IDP_t$

	BP/USD	DM/USD	JY/USD
$\Delta^2 p_t^*$	0.008 [0.244]	0.121 [1.472]	−0.047 [−4.049]
$\Delta^2 p_t$	−0.080 [−1.478]	0.204 [2.164]	−0.059 [−2.630]
Δi_t^*	0.006 [1.055]	0.001 [0.098]	0.002 [0.980]
Δi_t	0.025 [4.393]	−0.004 [−0.443]	−0.004 [−1.305]
Δgap_t	0.018 [2.564]	−0.050 [−2.995]	0.005 [1.990]
$\Delta(\widehat{s}_{t+1|t} - s_t)$	−0.617 [−3.719]	−2.728 [−4.474]	−0.122 [−1.841]
ΔIDP_t	0.156 [1.144]	−1.585 [−1.865]	2.541 [11.596]