Extracting Inflation from Stock Returns to Test Purchasing Power Parity

By BHAGWAN CHOWDHRY, RICHARD ROLL, AND YIHONG XIA*

Relative purchasing power parity (PPP) holds for pure price inflations, which affect prices of all goods and services by the same proportion, while leaving relative prices unchanged. Pure price inflations also affect nominal returns of all traded financial assets *by exactly the same amount. Recognizing that relative PPP may not hold for the official inflation data constructed from commodity price indices because of relative price changes and other frictions that cause prices to be "sticky," we provide a novel method for extracting a proxy for* realized *pure price inflation from stock returns. We find strong support for relative PPP in the short run using the extracted inflation measures.* (*JEL* F31, G15)

Purchasing power parity (PPP) is the simple proposition that prices in different countries should be equal if they are converted to the same currency. The absolute version of PPP is based on the law of one price, which maintains that arbitrage should tend to equilibrate prices of the same good at different locations. If the composition of the basket of goods used for constructing price indices is identical across countries, PPP trivially follows from the law of one price.

However, frictions to goods arbitrage such as transportation costs and other impediments to trade (the extreme being non-tradable goods such as land) inhibit cross-country price equalization. Even with such frictions, the relative version of PPP, which maintains that the *change* in price levels across countries should be the same after adjusting for the change in the exchange rate, may still hold if *relative* price changes across countries are identical. For instance, a pure money shock will change nominal prices of all goods, services, and assets by exactly the same amount so that relative prices among different assets will remain constant. In this strict pure price inflation case, the relative version of PPP will hold.

Although simple, the PPP hypothesis has defied empirical confirmation for decades. There seems to be little agreement about why it fails so spectacularly when taken to data.[1] The absolute version of PPP may fail because of the frictions to goods arbitrage mentioned above; this is not the focus of this paper. Our focus instead is on the failure of relative PPP.

The failure of relative PPP in its most basic form can be described as follows. If relative PPP holds, then changes in the exchange rate must equal the concurrent inflation differential between two countries. Empirically, the two are at most weakly correlated (Kenneth Rogoff, 1996). Furthermore, changes in exchange rates are extremely volatile, with a yearly standard deviation typically on the order of 12 to 13 percent for developed countries, while inflation

* Chowdhry: The Anderson School, UCLA, 110 Westwood Plaza, Los Angeles, CA 90095-1481 (e-mail: bhagwan@anderson.ucla.edu); Roll: The Anderson School, UCLA, 110 Westwood Plaza, Los Angeles, CA 90095-1481 (e-mail: rroll@anderson.ucla.edu); Xia: Department of Finance, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104 (e-mail: yxia@wharton.upenn.edu). The authors thank Werner Antweiler for providing the foreign exchange rate data; Kent Daniel, Ken French, and Xiaoyan Zhang for providing the Fama-French factor returns data; Ben Bernanke (the editor), three anonymous referees, Marshall Blume, Michael Brennan, Nai-fu Chen, John Cochrane, Bernard Dumas, Philippe Jorion, Andrew Karolyi, Yrjö Koskinen, seminar participants at the Tenth Georgia Tech International Finance Conference, California State University Fullerton, HEC France, the Stockholm School of Economics, and the brown-bag lunch seminar series at The Wharton School for useful comments on earlier drafts of the paper. Xia acknowledges the financial support of the Weiss Center at the Wharton School, University of Pennsylvania.

[1] Rather than provide a long list of relevant references here, we point the reader to an excellent survey by Rogoff (1996).

differentials have yearly standard deviations of less than 1 percent (see Rogoff, 1998). This empirical regularity has led scholars to wonder if exchange rate movements are "too" volatile to be explained by fundamentals. An alternative explanation consistent with the empirical regularity of high relative volatility of exchange rates is that prices of goods and services are "too sticky." Whether small menu costs or other transactions costs can create such a sluggish response in prices of goods and services has been debated in the literature theoretically and empirically (see Julio J. Rotemberg, 1982, N. Gregory Mankiw, 1985, Paul Krugman, 1986, Avinash K. Dixit, 1989a, 1989b, and Francisco A. Delgado, 1991). We make little contribution to this debate. Our contribution is in demonstrating that nominal price rigidities in the goods market may indeed be responsible for the failure of relative PPP, and once a proxy for the unobservable pure price inflation is found, relative PPP does hold, with the volatility of the inflation differentials being the same order of magnitude as the volatility of exchange rates.

We suspect that if consumption goods and services were freely and continuously traded in the same way as stocks, bonds, or foreign currencies, they would display price movements similar to those of financial assets and would be much more volatile than what a consumer observes in, say, a grocery store. But goods and services are not freely and continuously traded in financial markets, and their prices are sluggish in responding to monetary shocks (Rudiger Dornbusch, 1976), news, and consumer/investor expectations and sentiments. Moreover, if nominal prices were not "sticky," they would adjust not only to current monetary shocks and news or rumors in the market, but also to shocks that affect consumer and investor *expectations* of future price levels, as argued by Michael Mussa (1982). Mussa suggests that "information that changes these expectations can have a profound effect. . . . , even if the current observed change that embodies this information is seemingly not very large." Consistent with the arguments of David G. Barr and John Y. Campbell (1997), it seems possible that high-frequency variation in forward-looking price levels may be driven by "inflation scares"–scares that inflation may jump to very high levels or even spiral into hyperinflation. Inflation scares, however, may turn out to be justified only rarely and, as a result, the realized inflation series constructed with "sticky" commodity prices may not appear to be volatile. This is analogous to the "peso problem" in which spot prices of a pegged currency, interspersed with infrequent but large devaluations, may appear to be smooth for long periods; the high frequency fluctuations in forward prices of such a currency may be reflecting changing expectations of the size and possibility of devaluation. This suggests that the "true" unobservable pure price inflation is likely to be much more volatile at high frequencies than indicated by the official inflation measures derived from price indices.

Thus, one resolution of the PPP puzzle may come from using long horizon tests. This is extensively discussed in the literature and, indeed, PPP does tend to hold much better in the long run (Rogoff, 1996). An alternative strategy may be to obtain a high-frequency proxy for the unobservable pure price inflation. We wondered whether the PPP puzzle might be resolved in the short run as well by such a proxy, which would be free from problems of nominal rigidities, transaction costs, aggregation, and relative price changes often associated with official price measures such as the CPI. We attempt to answer this question by providing a novel method for extracting high-frequency "non-sticky" inflation proxies from stock returns.[2]

The essence of our method boils down to estimating the *nominal* return on a *real* risk-free asset. We label such a return the "ex post" nominal risk-free rate, because this rate includes both the real risk-free rate and the *realized* inflation rate during the period. Nominal returns on a traded risk-free asset would provide the most direct surrogate for our test of relative PPP in the short run. Unfortunately, there are no traded real risk-free assets available in the economy. Treasury bills are *nominal* risk-free assets, so their returns measure only nominal returns on nominal risk-free assets, which are risky in real terms due to the presence of *unexpected* inflation. Indexed securities such as Treasury Inflation Protected Securities (TIPS) are often considered real risk-free assets, but they are

[2] Stock and Watson (2003) provide an excellent survey of research on *forecasting* economic activity and inflation using asset prices.

available only for a limited time period and their returns are affected by contractual peculiarities.[3] One problem with U.S. TIPS, for example, is that their coupon payments are indexed to official inflation derived from the CPI, which is exactly the series we are trying to replace in our analysis.

In this study, we propose to extract the unobservable ex post nominal risk-free rate from nominal equity returns, which are readily available at high frequency and are very responsive to news and investor expectations. Nominal equity returns, however, must be purged of influences of real factors in order to get an estimate of the ex post nominal risk-free rate. A good asset pricing model will help minimize the residual real effects left in the extracted series, and we adopt the empirically successful Fama-French three-factor model (and an extension that includes momentum as a fourth factor) to describe the return generating process. An empirically–instead of theoretically–motivated model is chosen to ensure that none of the widely recognized and empirically successful common factors is omitted in purging the real effects from equity returns. A series extracted from the realized equity returns that is orthogonal to the common real factors will give a good measure of the ex post nominal risk-free rate. Our empirical procedure relies on the familiar Fama-MacBeth (1973) approach to extract such a series.

When the extracted proxies are used to replace the official CPI inflation measure in the PPP test, we find that relative PPP holds well in the short run in both single-country-pair OLS regressions and a pooled system regression. Results from our robustness checks suggest that our supportive evidence for short-run relative PPP is unlikely to be driven by missing world factors or by real effects of inflation. An examination of our extracted series and the CPI indicates long-run cointegration between the price level constructed from our series and the price level measured by the CPI, even though these two variables are only weakly correlated in the short run. This implies that the official CPI series may be a smoothed measure of inflation in the long run, while the inflation measure extracted from asset prices better captures temporary swings in inflation and inflation scares. Therefore, our results complement the findings from the long-run PPP tests and help resolve the PPP puzzle in the short run.

The paper is organized as follows: Section I introduces the theoretical framework and presents a detailed derivation of the empirical methodology; Section II describes the data; Section III presents empirical results and carries out robustness checks; Section IV summarizes and concludes the paper.

I. Theoretical Framework and Empirical Methodology

In this section, we first present the factor model for stock returns and derive the empirical method for extracting the ex post nominal risk-free rate from stock returns. In the second subsection, we report the PPP test using our extracted inflation series.

A. *Extracting the Nominal Risk-Free Rate from Stock Returns*

To study PPP, we require estimates of the pure price inflation, which is linearly related to the nominal return on a *real* risk-free asset:

$$R_{ft} \equiv r_{ft} + \pi_t$$

where r_{ft} denotes the realized real risk-free rate and π_t denotes the realized pure price inflation for the period from date $t - 1$ to t. The most obvious way to obtain R_{ft} is to use the nominal return on a traded risk-free asset, and the widely used Treasury bill rates come to mind. Unfortunately, Treasury bill rates are the nominal rates on *nominal* risk-free assets. Treasury bill rates for period from date $t - 1$ to t, TB_{t-1}, are determined at date $t - 1$ (at the beginning of the period) based on investors' expectations of the real rates and the inflation for period from date $t - 1$ to t. What happens if there is a 1-percent unexpected pure price inflation during this period? The rate TB_{t-1} is not affected because it is fixed at date $t - 1$, but the nominal return on a real risk-free asset R_{ft}, which is realized at the end of the period, goes up by 1 percent since it

[3] The U.S. TIPS started trading only in January 1997. The United Kingdom has a longer history of trading TIPS, but they are not very well indexed. The lag is eight months between CPI changes and the revision in the coupon payment of U.K. TIPS; this induces a lot of nominal risk and pollutes the real yields.

is indexed by the realized pure price inflation during the period from $t - 1$ to t. Therefore, R_{ft} instead of TB_{t-1} is the correct variable to use in testing relative PPP relation.

As we mentioned earlier, one should not use the realized nominal returns on securities such as TIPS as a measure of R_{ft} either, because the coupons on these securities are indexed to official inflation data such as the changes in the CPI, which we argue is not a good measure of pure price inflation.

The alternative is to extract R_{ft} from observable nominal returns on risky financial assets. Financial asset prices have some advantages over commodity prices. First, financial asset prices capture all relevant information and summarize investors' expectations or perceptions of the economy including inflation. Moreover, financial asset prices are readily available without any time lag or any problems of aggregation. Therefore, pure price inflation extracted from asset returns, as compared to official inflation data, is likely to be more responsive to "news" or "perceptions" in financial markets and is relatively unaffected by frictions such as transportation costs and nominal rigidities.

Since risky asset returns are influenced by many real factors, the ex post nominal risk-free returns can be extracted only after we purge nominal asset returns of influences of real factors. This requires an asset pricing model. Instead of developing an asset pricing model from fundamentals, we adopt the practical view that an empirically successful model can capture, at least in sample, most real effects in asset returns, allowing us to extract a relatively good estimate of the ex post nominal risk-free return.

We adopt the popular Fama-French three-factor model (and an extension that includes momentum as a fourth factor) to describe the excess real return generating process:

$$r_{it} - r_{ft} = \sum_{k=1}^{3} \beta_{ik} f_{kt} + \epsilon_{it} \tag{1}$$

where r_{it} is the real return on asset i for the period from $t - 1$ to t, r_{ft} is the real risk-free rate, and the three factors f_{kt} are approximated by Fama-French three factors: (1) returns on the market index in excess of the risk-free rate, $r_{Mt} - r_{ft}$; (2) returns on the zero-investment SMB portfolio, $r_{St} - r_{Bt}$, where r_{St} (r_{Bt}) is the return on a small (big) cap portfolio; and (3) returns on the zero-investment HML portfolio, r_{Ht}-r_{Lt}, where r_{Ht} (r_{Lt}) is the return on a high (low) book-to-market portfolio. In addition, ϵ_{it} is a spherical disturbance and the β's are constant factor loadings.

Equation (1) applies to real returns, but we observe only nominal returns for financial assets. To transform the real return-generating process to nominal returns, we further assume that the Fisher equation holds so that

$$R_{it} = r_{it} + \pi_t, \ \forall i$$

where r_{it} denotes the real return on asset i, R_{it} denotes its nominal correspondent, while π_t stands for the *pure price inflation*. Note that the expected and unexpected inflation may have real effects on asset returns and the real effect may be different for different assets.[4] To derive a tractable empirical model, we assume that all potential real effects, including the real effects of inflation, are captured by the three factors. Under this assumption, r_{it} captures all the real effects while π_t measures only *pure price inflation* that is free from any real effects and is thus orthogonal to r_{it} ($\forall i$). By definition, π_t has a one-to-one relation with all nominal returns R_{it} ($\forall i$) and does not induce differential changes in nominal returns. To illustrate this idea, consider a real event such as the assassination of Mexican presidential candidate Donaldo Colosio in March 1994. The peso depreciated against the U.S. dollar when the assassination took place. The Mexican peso prices of arbitrageable assets, such as gold, rose as a result of the exchange rate depreciation but, at the same time, Mexican peso relative prices of other assets must have declined because there was no immediate pure inflation, π_t. The depreciation of the Mexican peso in this case was real. In contrast, if the Mexican government decided to increase the money supply, all nominal returns would increase and the common component in returns attributable to π_t would be significant. The Mexican peso would depreciate in this case

[4] Jacob Boudoukh et al. (1994) argue that expected inflation may be correlated with future production and thus affect real stock returns. We will later check the impact of adding expected inflation as an additional factor in the model.

as well but there would be little, if any, depreciation of the currency in real terms.

In comparison, the official CPI inflation is calculated from a basket of consumption goods. For the purpose of testing PPP, CPI inflation is contaminated by the movement in relative prices across different goods in the basket, the change in the composition of the basket, and the presence of non-traded commodities.

Under the assumption of $R_{it} = r_{it} + \pi_t$, the return-generating process given in equation (1) can be formulated in terms of observable nominal returns

$$(2) \qquad R_{it} - R_{ft} = \beta_{i1}[R_{Mt} - R_{ft}] + \beta_{i2}[R_{St} - R_{Bt}] + \beta_{i3}[R_{Ht} - R_{Lt}] + \epsilon_{it}$$

where $R_{ft} \equiv r_{ft} + \pi_t$ is the unobservable ex post nominal risk-free return. To estimate R_{ft}, we employ the standard two-stage Fama-MacBeth (1973) regression methodology.

In the first stage, a time series regression is carried out to estimate the betas. To capture any possible mispricing of the model, an intercept α_i is introduced to (2), which is then rewritten as follows:

$$(3) \quad R_{it} - \mathrm{TB}_{t-1} = \alpha_i + \beta_{i1}[R_{Mt} - \mathrm{TB}_{t-1}] + \beta_{i2}[R_{St} - R_{Bt}] + \beta_{i3}[R_{Ht} - R_{Lt}] + \eta_{it}$$

where

$$\eta_{it} \equiv (1 - \beta_{i1})[R_{ft} - \mathrm{TB}_{t-1}] + \epsilon_{it}.$$

Since $\mathrm{TB}_{t-1} = \mathrm{E}_{t-1}[R_{ft}]$, $R_{ft} - \mathrm{TB}_{t-1}$ measures the unexpected inflation (plus the real rate) realized during the period. The error term η_{it} is comprised of two mean zero terms: the idiosyncratic risk, ϵ_{it}, and a linear function of the market-wide unexpected inflation (plus the real rate), $(1 - \beta_{i1})[R_{ft} - \mathrm{TB}_{t-1}]$.

The second stage of estimation consists of a cross-sectional regression on each date t. Rewriting equation (1), and including the mispricing term $\hat{\alpha}_i$:

$$(4) \quad R_{it} - \hat{\alpha}_i = \hat{R}_{ft} + \sum_{k=1}^{3} f_{kt}\hat{\beta}_{ik} + \epsilon_{it} \ \ \forall i.$$

The intercept of the above cross-sectional regression of stock returns (minus the possible mispricing) on the three-factor betas provides the estimate of R_{ft}. While the traditional Fama-MacBeth approach yields a time series estimate of $\hat{R}_{ft}$, most researchers are interested only in its sample mean $\bar{R}_{ft} \equiv 1/T \sum_{t=1}^{T} \hat{R}_{ft}$. We, however, are interested primarily in the time series behavior of $\hat{R}_{ft}$.

In our first-stage time-series regression, the nominal returns in excess of the Treasury bill rate are regressed on the excess market portfolio returns and the two Fama-French factor returns. To mitigate estimation errors in the betas, we employ industry portfolios for both domestic and foreign stocks as the base assets. This step produces estimates of the betas. Notice that because the error term η_{it} in (3) includes unexpected inflation, it may be correlated with the realized nominal return on the market and consequently bias the estimate of β's.[5]

The Fama-French factors are created using domestic stocks only. One might argue that with integrated world capital markets, return generating factors might include some world factors as well.[6] If world factors were included,

[5] In general, the fact that β's are estimated in the first stage and then used in the second stage as regressors introduces the errors-in-variable problem in the second-stage regression. Although a single-step regression can avoid this problem, it is infeasible in our current setting, where we are interested in the whole time series estimates of $\hat{R}_{ft}$. Adrian Pagan (1984) shows, in the setting of using estimated variables as regressors, that the two-step estimator is consistent and asymptotically efficient and there is no efficiency gain by switching to a full MLE, despite the result that the parameter variances are estimated inconsistently. Jay Shanken (1992) provides standard error corrections for the second-stage parameter estimates. Since we are interested only in the point estimates, a correction for the standard errors is not relevant here.

[6] For example, Andrew Ang and Geert Bekaert (2001) find that the U.S. rate is a stronger instrument than local ones in predictive regressions when the local excess returns are converted into U.S. returns, lending support for a globally integrated market. We later check the impact of adding the U.S. short rate as an additional factor in our model and find that our PPP results are virtually unchanged. Xiaoyan Zhang (2005) also finds that the international CAPM with foreign exchange risk performs better than the single beta international CAPM and the Fama-French international

however, we would have to convert returns denominated in foreign currencies into domestic units, which then introduces a foreign exchange rate component in our estimates. We want to avoid having exchange rates appear in both the dependent and independent variables in our PPP tests. Furthermore, John M. Griffin (2002) finds that domestic Fama-French factors explain much more time-series variation in returns and generally have lower pricing errors than the (Fama-French) world factor model.

Note that the ex post nominal risk-free rate differential, $R_{ft}^{*} - R_{ft}$, includes both pure price inflation differential, $\pi_t^{*} - \pi_t$, and the real interest-rate differential, $r_{ft}^{*} - r_{ft}$, where foreign variables are denoted with superscript *. Since the real interest rate differential is unobservable, it is only feasible for us to use the ex post nominal risk-free rate differential to approximate for the inflation differential. This approximation is harmless for the purpose of testing PPP relation if we put the noisy approximation on the left-hand side of the regression and if the real interest rate differential $r_{ft}^{*} - r_{ft}$ is correlated with neither the pure price inflation differential nor the foreign exchange rate changes.[7] From a theoretical point of view, many general equilibrium models often assume that the *real* interest rate in the economy is determined by such fundamentals as the productivity of the economy and is uncorrelated with either realized pure price inflation or foreign exchange rates. Empirically speaking, although several recent papers find a positive relationship between real interest rate differentials and the spot rate changes over the long run (three to five years),[8] there is virtually no covariation between the two variables with a horizon of under one year. Since we are interested in examining the PPP in the short run at the monthly, bimonthly, and quarterly horizons, the contamination of the realized inflation measure by the realized real interest rate differentials will be unlikely to bias our estimate, although it will reduce the regression R^2 and affect the calculation of the standard error.

B. *Testing PPP*

The final step is the test of the PPP hypothesis. The relative PPP hypothesis implies that

$$\Delta s_t = \Delta \pi_t \equiv \pi_t^{*} - \pi_t \tag{5}$$

where s_t denotes the logarithm of the nominal exchange rate defined as the price of domestic currency in units of foreign currency (e.g., yen per dollar) and π_t^{*} and π_t denote the change in log price indices and thus are the inflation rates in foreign and domestic countries, respectively.

Although the theory itself does not specify whether $\Delta\pi$ or Δs should be the dependent variable, changes in spot rates are usually regressed on the inflation differentials in the empirical test of relative PPP:

$$\Delta s_t = \delta + \gamma \Delta \pi_t + \epsilon_t \tag{6}$$

where $\delta + \epsilon$ captures the movement of spot rate changes unaccounted for by inflation differentials. Under the null hypothesis that relative PPP holds, we have $H_0 : \delta = 0$ and $\gamma = 1$. The alternative $\delta \neq 0$ suggests that exchange rates, on average, move for other reasons as well, which contributes to the higher volatility of exchange rate changes, while the alternative $\gamma > 1$ incorporates the possibility of exchange rate overshooting models (Dornbusch, 1976).

It seems natural to put Δs as independent variables simply because it is much more volatile than the problematic official CPI inflation differentials. The implicit assumption behind this conventional regression specification seems to imply that measurement errors or noises in foreign exchange rate changes are more severe than those in the inflation differentials, even though the exchange rates are directly determined in the market and should be more reliable measures than official indices.

Unlike the conventional specification, we test the relative PPP hypothesis by regressing $\Delta\pi$ on Δs. This is because Δs is directly measured in the market while we can only at best use $\Delta\hat{R}_{ft} \equiv \hat{R}_{ft}^{*} - \hat{R}_{ft}$ as a surrogate for $\Delta\pi$ and, as explained in the previous subsection, $\hat{R}_{ft}$ contains substan-

three-factor model. We later check whether our time series regression residuals are related to a world factor proxied by the Morgan Stanley Composite Index (MSCI), which is almost perfectly correlated with the Datastream Global Index used by Zhang (2005).

[7] For example, Michael Adler and Bruce Lehman (1983) assume that real interest rate differential is constant and thus satisfies both conditions.

[8] See, for example, Hali J. Edison and B. Dianne Pauls (1993), Marianne Baxter (1994), and Michael Bleaney and Douglas Laxton (2003).

tial noise. Our regression specification is thus given by

$$\Delta \hat{R}_{ft} = a + b\Delta s_t + \epsilon_t \tag{7}$$

which is the observable counterpart to $\Delta\pi_t + (r^*_{ft} - r_{ft}) = a + b\ \Delta s_t + \epsilon_t$. Under the null hypothesis that PPP holds so that $\sigma(\Delta\pi_t) \approx \sigma(\Delta s_t)$, $b = 1$. Moreover, the intercept of the regression is given by the mean real interest rate differential. If the average real interest rates in foreign and domestic countries are close to each other during the sample period, then we would expect $a = 0$ under the null.

Our null hypothesis is that PPP holds if we have a good measure of the unobservable pure price inflation. An obvious alternative argument, widely used in the literature, is that inflation is measured correctly by the official series, but relative PPP fails for other reasons. Two well-known alternatives are exchange rate overshooting so that $\gamma > 1$ and excess exchange rate volatility unrelated to inflation differential so that $\sigma^2(\Delta s_t) > \sigma^2(\Delta\pi_t)$. Under both these alternatives, the slope coefficient would be close to zero.

The specification of (7) could be easily confused with the specification of the Uncovered Interest Rate Parity (UCIRP) relation, which states that the *expected* change in the exchange rates equals the nominal interest rate differentials.[9] In our notation, this is equivalent to testing the following relation:

$$E_{t-1}(\Delta s_t) = TB^*_{t-1} - TB_{t-1}$$

where the foreign and domestic Treasury bill rates, TB^*_{t-1} and TB_{t-1}, are predetermined at the beginning of period t.

In empirical studies, the actual change of spot rate in period t, Δs_t, is often used as a proxy for the expected change, $E_{t-1}(\Delta s_t)$. In contrast, Δs_t is not a proxy for any expected variables in our specification. What we are after is the relation between the realized spot rate change and the realized nominal return of a real risk-free (or perfectly indexed) asset.

[9] Numerous empirical studies find that uncovered interest rate parity does not hold, leading to the conjecture that foreign exchange rate risks are priced.

II. The Data

We use three sets of data. The first includes industry stock returns and the three Fama-French factors, namely, the excess market return, the return on a zero-investment portfolio of SMB,[10] and the return on a zero-investment portfolio of HML,[11] in the United States, the United Kingdom, Japan, and Germany. The second set of data includes exchange rates defined as the foreign currency (yen, British pound, and German mark) per U.S. dollar. The third set of data includes government CPI inflation measures.

All the U.S. industry returns, the Treasury bill rates, and the three Fama-French factors are from Kenneth French's Web site,[12] with the sample ranging from July 1926 to December 2000. The U.K., German, and Japanese industry returns are from the Datastream Global Index. For the United Kingdom and Germany, total industry returns including dividends are available, while for Japan only capital gains industry returns are available. The market returns for the United Kingdom, Japan, and Germany are constructed using the total (including dividend) market returns from Datastream, and the SMB and HML factor returns are constructed using raw data from Datastream as well.[13] While the Treasury bill rates for the United Kingdom and Japan are from Datastream, the rate for Germany is from Bloomberg. The sample period for the United Kingdom is from January 1986 to December 1999 (168 monthly observations); for Japan it is from May 1983 to December 1999 (200 monthly observations); and for Germany it is from January 1988 to December 1999 (144 monthly observations).

The change in the foreign exchange rate is calculated from the end of month to the end of month using the daily foreign exchange rate provided by Pacific Foreign Exchange Rate Service.[14] We also calculated the foreign exchange rate changes from the beginning of the month to

[10] Long small firms and short big firms.

[11] Long high book-to-market firms and short low book-to-market firms.

[12] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/.

[13] Zhang and Kent Daniel kindly provided the SMB and HML data to us.

[14] The foreign exchange rate is from http://pacific.commerce.ubc.ca/xr/.

TABLE 1—DATA SUMMARY STATISTICS

Variable	Mean	Standard deviation
US market	1.36	4.24
US SMB	−0.20	2.67
US HML	0.15	2.72
US Tbill	0.48	0.16
US CPI inflation	0.27	0.20
US extracted risk-free rate	0.48	7.15
UK market	1.39	4.83
UK SMB	0.05	4.11
UK HML	0.26	2.55
UK Tbill	0.71	0.26
UK £-US $	−0.07	3.14
UK CPI inflation	0.33	0.47
UK extracted risk-free rate	0.71	8.22
German market	1.34	5.15
German SMB	−0.44	4.41
German HML	0.28	3.26
German Tbill	0.44	0.17
German DM-US $	0.15	3.10
German CPI inflation	0.19	0.26
German extracted risk-free rate	0.44	7.37
Japan market	0.75	5.92
Japan SMB	0.03	3.34
Japan HML	0.13	3.07
Japan Tbill	0.30	0.21
Japanese ¥-US $	−0.42	3.52
Japanese CPI inflation	0.10	0.47
Japanese extracted risk-free rate	0.30	8.73

Notes: This table reports sample mean and sample volatility for the three Fama-French factor returns for the United States, the United Kingdom, Germany, and Japan. The sample mean and volatility for the change in the foreign exchange rate are also reported. The sample period for the United States and Japan is from May 1983 to December 1999, for the United Kingdom is from January 1986 to December 1999, and for Germany is from January 1988 to December 1999. The numbers are percentage per month.

the beginning of the month, and the empirical results were virtually unchanged.

Table 1 provides summary statistics for these three datasets. The market excess return for the United States, United Kingdom, and Germany is around 1.3 to 1.4 percent per month with a sample standard deviation around 4 to 5 percent per month, but the Japanese market return has much lower mean and slightly higher volatility. The SMB and HML factors have much smaller mean returns. One-month Treasury bill rates vary from a high of 0.71 percent per month for the United Kingdom to a low of 30 basis points in Japan. Official measures of inflation are calculated from CPI data, and monthly average rates are around 0.27 percent, 0.33 percent, 0.19 percent, and 0.10 percent for the United States, United Kingdom, Germany, and Japan, respectively.

As is well known in the literature, there is a striking contrast among the volatilities of Treasury bill rates, CPI inflation rates, and spot rate changes. For example, the sample volatilities of Treasury bill rates are smaller than 30 basis points and CPI inflation rates have sample volatilities smaller than 50 basis points for all four countries. In contrast, spot exchange rate changes for the three currency pairs are all above 3 percent per month. This is not surprising if spot rates move not only with actual realized inflation but also with "inflation scares," which could be driven by high-frequency economic rumors that materialize only infrequently and thus rarely affect official CPI inflation measures.

III. Empirical Results

Our preliminary empirical results for individual country pairs (GM-US, JP-US, and UK-US), first based on the Fama-French three-factor model and then based on the Carhart (1997) four-factor model, are presented in the first two subsections. In the third subsection, we pool the three country pairs in a system regression, which provides a unified framework across all three country pairs and enhances test power as well. In the last subsection, we provide several robustness checks and also examine the relation between our equity-extracted inflation and the official CPI inflation.

A. *Results under the Three-Factor Model*

In this subsection, we provide empirical evidence for the relative PPP hypothesis using inflation extracted from equity returns.

First, a time series regression is carried out based on equation (3):

$$R_{it} - \mathrm{TB}_{t-1} = \alpha_i + \beta_{i1}(R_{Mt} - \mathrm{TB}_{t-1}) + \beta_{i2}R_{\mathrm{SMB},t} + \beta_{i3}R_{\mathrm{HML},t} + \eta_{it}$$

where R_{it} is industry i's portfolio return in period t measured in local currency and R_{Mt}, $R_{\mathrm{SMB},t}$, and $R_{\mathrm{HML},t}$ are concurrent country-specific factors also measured in local currency.

TABLE 2—TIME SERIES REGRESSIONS OF EXCESS INDUSTRY RETURNS ON THE THREE FAMA-FRENCH FACTORS (UNITED STATES)

Industry	Constant		R_m − *Tbill*		R_{SMB}		R_{HML}		Adj. R^2
Food	−0.017	(0.27)	0.967	(0.06)	−0.444	(0.12)	0.044	(0.17)	0.633
Beer	0.103	(0.27)	0.798	(0.08)	−0.360	(0.11)	−0.268	(0.16)	0.500
Smoke	0.176	(0.54)	0.926	(0.11)	−0.195	(0.23)	0.154	(0.33)	0.271
Games	0.099	(0.31)	1.174	(0.07)	0.336	(0.17)	−0.103	(0.22)	0.699
Books	−0.216	(0.19)	1.088	(0.06)	0.087	(0.07)	0.196	(0.12)	0.757
Household	0.198	(0.14)	1.018	(0.04)	−0.389	(0.06)	−0.144	(0.07)	0.851
Apparel	−0.872	(0.33)	1.180	(0.08)	0.462	(0.12)	0.232	(0.22)	0.648
Health	0.247	(0.24)	0.902	(0.06)	−0.404	(0.09)	−0.456	(0.18)	0.710
Chemicals	−0.164	(0.20)	1.129	(0.06)	0.051	(0.07)	0.424	(0.17)	0.685
Textiles	−0.582	(0.36)	1.116	(0.08)	0.854	(0.12)	0.681	(0.18)	0.615
Construction	−0.324	(0.20)	1.192	(0.04)	0.245	(0.07)	0.307	(0.09)	0.825
Steel	−0.355	(0.30)	1.149	(0.07)	0.533	(0.11)	0.407	(0.15)	0.610
Fabricated products	−0.291	(0.25)	1.163	(0.04)	0.626	(0.08)	0.286	(0.10)	0.776
Electronic equipment	0.195	(0.34)	1.076	(0.06)	0.538	(0.12)	−0.501	(0.25)	0.761
Autos	−0.296	(0.26)	1.262	(0.05)	0.216	(0.08)	0.716	(0.13)	0.708
Carry	−0.300	(0.28)	1.147	(0.09)	0.081	(0.10)	0.339	(0.17)	0.640
Mines	−0.355	(0.43)	0.795	(0.11)	0.790	(0.16)	0.371	(0.19)	0.308
Coal	−0.712	(0.36)	0.974	(0.08)	0.634	(0.16)	0.559	(0.20)	0.455
Oil	−0.035	(0.28)	0.854	(0.07)	0.051	(0.13)	0.540	(0.12)	0.447
Utilities	−0.165	(0.22)	0.611	(0.06)	−0.284	(0.11)	0.580	(0.14)	0.462
Telecommunication	0.437	(0.29)	0.928	(0.07)	−0.156	(0.09)	−0.022	(0.13)	0.627
Services	0.638	(0.18)	1.004	(0.05)	0.262	(0.10)	−0.882	(0.13)	0.848
Business equipment	0.266	(0.31)	0.985	(0.07)	0.305	(0.14)	−0.565	(0.19)	0.696
Paper	−0.273	(0.18)	1.088	(0.06)	0.105	(0.09)	0.267	(0.16)	0.682
Trans	−0.507	(0.24)	1.162	(0.06)	0.325	(0.10)	0.496	(0.10)	0.702
Wholesale	−0.317	(0.17)	1.025	(0.05)	0.426	(0.06)	0.016	(0.11)	0.870
Retail	0.200	(0.23)	1.085	(0.07)	0.163	(0.08)	−0.168	(0.10)	0.746
Meals	−0.302	(0.24)	1.079	(0.06)	0.284	(0.12)	0.091	(0.14)	0.735
Finance	−0.232	(0.14)	1.159	(0.05)	−0.160	(0.07)	0.474	(0.07)	0.881
Other	−0.775	(0.32)	1.199	(0.06)	0.258	(0.09)	0.265	(0.21)	0.696

Notes: This table reports the regression results of excess industry portfolio returns on the three Fama-French factors for the United States with sample period from January 1986 to December 1999. The Newey-West adjusted standard errors are reported in parentheses to the right of the coefficient estimates.

This procedure is repeated for the United States, United Kingdom, Japan, and Germany.

Time-series regressions of the industry portfolio returns are reported in Tables 2 to 5. The adjusted R^2 varies materially across industries, from as low as 17 percent to as high as 88 percent.[15] To account for the possibility that the residual term η may be serially correlated and heteroskedastic, we always report the Newey-West (1987) adjusted standard errors.

Next, cross-sectional regression (4) is computed, with industry returns (minus estimated mispricing) regressed on the beta estimates from the time series regressions. We assume that the cross-sectional residual vector $\epsilon_t \equiv [\epsilon_{1t}, \ldots, \epsilon_{nt}]$ at time t is uncorrelated with the portfolio's beta estimates, $\hat{\boldsymbol{\beta}}_k \equiv [\hat{\beta}_{k1}, \ldots, \hat{\beta}_{kn}]$ $(k = 1,2,3)$, so the estimate of R_{ft} at each period t is unbiased except for the impact from estimation errors in beta.[16]

The estimates for R_{ft}^* and R_{ft} are stored for each t. We find that both series exhibit extraordinarily high sample volatility, which is probably due to the estimation error in both the time series and the cross-sectional regressions. The sample mean and volatility are reported in

[15] The widely varying R^2 in the time series regression is not an issue so long as it reflects large idiosyncratic noise in some industry portfolio returns, instead of systematic risk remaining in the residual.

[16] Our simulation evidence, not reported for brevity, suggests that our procedure is, in general, quite effective in extracting R_{ft}. The second-stage estimate $\hat{R}_{ft}$ is generally unbiased, but $\hat{R}_{ft}$ may be measured with considerable noise. Please also refer to footnote 5 for more discussion of this issue.

TABLE 3—TIME SERIES REGRESSIONS OF EXCESS INDUSTRIAL RETURNS ON THE THREE FAMA-FRENCH FACTORS (UNITED KINGDOM)

Industry	Constant		$R_m - Tbill$		R_{SMB}		R_{HML}		Adj. R^2
Mining	0.166	(0.42)	1.183	(0.12)	0.212	(0.20)	0.187	(0.27)	0.426
Oil & gas	0.229	(0.28)	0.878	(0.09)	−0.028	(0.08)	0.507	(0.18)	0.518
Chemicals	−0.316	(0.32)	1.060	(0.06)	0.111	(0.09)	−0.168	(0.18)	0.631
Construction	−0.749	(0.28)	1.270	(0.06)	0.141	(0.08)	0.543	(0.13)	0.718
Forestry & paper	−0.024	(0.54)	1.171	(0.13)	0.392	(0.19)	0.073	(0.30)	0.303
Steel	−0.008	(0.45)	1.300	(0.10)	0.308	(0.15)	0.507	(0.21)	0.402
Aerospace & defense	−0.377	(0.38)	1.117	(0.10)	0.164	(0.16)	−0.046	(0.22)	0.526
Diversified industrials	−0.610	(0.27)	1.076	(0.05)	0.050	(0.06)	−0.042	(0.15)	0.660
Electronic equipment	0.240	(0.35)	1.091	(0.13)	0.449	(0.24)	−0.497	(0.27)	0.529
Engineering machinery	−0.346	(0.28)	1.223	(0.07)	0.291	(0.09)	−0.005	(0.17)	0.691
Autos	0.084	(0.38)	1.256	(0.09)	0.142	(0.11)	−0.106	(0.19)	0.568
Textiles	−1.260	(0.41)	1.202	(0.09)	0.284	(0.14)	0.222	(0.22)	0.640
Beverages	−0.128	(0.25)	0.928	(0.07)	−0.226	(0.11)	−0.242	(0.14)	0.685
Food	−0.183	(0.23)	0.810	(0.07)	−0.047	(0.11)	−0.092	(0.17)	0.651
Health	−0.294	(0.21)	0.943	(0.05)	0.183	(0.06)	−0.348	(0.16)	0.670
Pack	−0.347	(0.34)	1.128	(0.07)	0.178	(0.12)	−0.070	(0.17)	0.558
Personal care	−0.323	(0.36)	0.826	(0.08)	−0.203	(0.10)	0.014	(0.19)	0.389
Pharmaceutical & biotechnology	0.782	(0.30)	0.840	(0.08)	−0.097	(0.09)	−0.901	(0.17)	0.573
Tobacco	0.346	(0.47)	0.776	(0.16)	−0.308	(0.17)	0.214	(0.25)	0.326
Dist.	−0.409	(0.28)	1.182	(0.06)	0.307	(0.08)	0.015	(0.11)	0.693
General retailers	−0.516	(0.24)	0.904	(0.05)	−0.174	(0.08)	0.209	(0.14)	0.664
Entertainment & hotels	−0.312	(0.28)	1.178	(0.08)	0.075	(0.08)	0.316	(0.11)	0.743
Media	0.064	(0.22)	1.272	(0.06)	0.371	(0.10)	−0.184	(0.10)	0.758
Restaurants & pubs	−0.144	(0.26)	0.819	(0.08)	−0.154	(0.06)	0.269	(0.11)	0.620
Support	−0.096	(0.20)	1.085	(0.05)	0.375	(0.07)	−0.138	(0.08)	0.725
Transport	−0.261	(0.25)	1.023	(0.05)	0.058	(0.07)	0.200	(0.15)	0.732
Food & drug retailers	−0.225	(0.33)	0.684	(0.06)	−0.126	(0.09)	0.186	(0.18)	0.385
Telecom services	0.605	(0.38)	0.856	(0.09)	−0.134	(0.11)	−0.200	(0.17)	0.498
Banks	0.440	(0.25)	1.117	(0.08)	−0.346	(0.08)	0.436	(0.19)	0.720
Insurance	−0.387	(0.25)	1.073	(0.07)	−0.166	(0.08)	0.181	(0.11)	0.679
Life assurance	0.479	(0.28)	0.925	(0.06)	−0.122	(0.07)	0.067	(0.14)	0.557
Investment firms	−0.100	(0.15)	1.066	(0.05)	0.165	(0.06)	0.031	(0.07)	0.870
Real estate	−0.761	(0.23)	1.003	(0.08)	0.109	(0.07)	0.943	(0.13)	0.718
Other financials	−0.054	(0.25)	1.298	(0.07)	0.297	(0.07)	0.106	(0.18)	0.748
IT hardware	1.367	(1.14)	1.412	(0.27)	0.845	(0.49)	−0.904	(0.82)	0.253
Software	0.649	(0.47)	1.136	(0.11)	0.857	(0.22)	−0.461	(0.24)	0.505

Notes: This table reports the regression results of excess industry portfolio returns on the three Fama-French factors for the United Kingdom with sample period from January 1986 to December 1999. The Newey-West adjusted standard errors are reported in parentheses to the right of the coefficient estimates.

Table 1. The sample means of $\hat{R}_{ft}$ of the four countries, by construction, are exactly the same as those of the Treasury bill rates. The sample volatility, however, is above 7 percent per month, which is a very large number even in comparison with the volatility of spot exchange rate changes. This reflects large measurement errors in our estimates of R_{ft} and is the main reason that we put the estimated nominal rate differentials on the left-hand side of the PPP regression as discussed below.

For further comparison, in Table 6 we report summary statistics for the extracted risk-free rate differentials, $\Delta\hat{R}_{ft} \equiv \hat{R}^*_{ft} - \hat{R}_{ft}$, official CPI inflation differentials, $\Delta\pi_{\mathrm{CPI}} \equiv \pi^*_{\mathrm{CPI}} - \pi_{\mathrm{CPI}}$, and spot exchange rate changes, Δs. The table also provides correlation among the three variables.

As is well known in the literature, the volatilities of official CPI inflation differentials are *too small* to be comparable to those of the exchange rate changes: the former are only around one-fifth to around one-tenth of the latter. On the other hand, our extracted risk-free rate is uniformly more volatile than the exchange rate changes, with the extracted risk-free rate differential volatilities around two to three times the exchange rate change volatilities.

TABLE 4—TIME SERIES REGRESSIONS OF EXCESS INDUSTRIAL RETURNS ON THE THREE FAMA-FRENCH FACTORS (GERMANY)

Industry	Constant		$R_m - Tbill$		R_{SMB}		R_{HML}		Adj. R^2
Autos	−0.343	(0.34)	1.118	(0.09)	−0.287	(0.14)	0.179	(0.19)	0.736
Banks	−0.220	(0.23)	0.911	(0.07)	−0.255	(0.08)	0.378	(0.12)	0.782
Chemicals	−0.058	(0.25)	0.825	(0.08)	−0.239	(0.09)	0.308	(0.10)	0.698
Media	1.126	(0.67)	0.853	(0.14)	0.454	(0.16)	−0.656	(0.26)	0.229
BSC resources	0.082	(0.27)	1.120	(0.08)	0.463	(0.09)	0.223	(0.12)	0.663
Food & beverages	−0.205	(0.17)	0.851	(0.06)	0.612	(0.07)	−0.034	(0.07)	0.618
Technology	−0.010	(0.31)	1.081	(0.07)	−0.090	(0.09)	0.018	(0.11)	0.698
Insurance	0.075	(0.29)	1.010	(0.08)	−0.167	(0.11)	−0.234	(0.17)	0.712
Transportation & logistics	−0.048	(0.43)	1.190	(0.13)	0.280	(0.17)	0.305	(0.13)	0.463
Machinery	−0.554	(0.25)	1.244	(0.06)	0.481	(0.08)	0.215	(0.09)	0.753
Industrial	1.252	(0.43)	1.071	(0.11)	−0.010	(0.11)	−0.251	(0.24)	0.584
Construction	−0.483	(0.39)	1.475	(0.10)	1.003	(0.17)	0.036	(0.11)	0.586
Pharmaceutical & health	0.440	(0.26)	0.878	(0.06)	0.274	(0.11)	−0.261	(0.12)	0.555
Retail	−0.376	(0.37)	1.247	(0.13)	0.593	(0.10)	−0.185	(0.14)	0.546
Software	2.578	(0.83)	1.434	(0.15)	0.467	(0.22)	−0.087	(0.25)	0.308
Telecom	0.284	(0.67)	0.738	(0.13)	0.076	(0.18)	−0.397	(0.27)	0.170
Utilities	0.063	(0.23)	0.689	(0.07)	−0.009	(0.09)	0.222	(0.08)	0.594
Financial services	0.552	(0.33)	0.935	(0.13)	0.518	(0.14)	−0.123	(0.12)	0.408
Consumer cyclical	−0.541	(0.40)	1.188	(0.08)	0.604	(0.11)	0.082	(0.14)	0.545

Notes: This table reports the regression results of excess industry portfolio returns on the three Fama-French factors for Germany with sample period from January 1988 to December 1999. The Newey-West adjusted standard errors are reported in parentheses to the right of the coefficient estimates.

These larger extracted risk-free rate differential volatilities, however, all lie in plausible range, since they should be at least as large as the exchange rate change volatilities if PPP is valid and if, in addition, they contain sizable estimation noise. As expected, our extracted risk-free rate differentials are more volatile than the official inflation differentials. This is consistent with the influence of changing expectations absent from the official series, and thus resolves the tension between the long-observed grossly larger volatility of exchange rate changes as compared to official inflation differentials.

Despite dramatic differences in standard deviations, the correlations between our extracted and the official inflation differentials are mostly larger than 0.1. The correlations between the official series and the exchange rate changes are positive for the Germany-U.S. pair, but are negative for the Japan-U.S. and the U.K.-U.S. pairs. In contrast, the correlations between our extracted series and the spot rate changes are all positive and uniformly larger: the correlation is as high as 0.5 for quarterly Germany-U.S. data.

Finally, we use $\Delta\hat{R}_{ft}$ as a surrogate for $\Delta\pi$ to test the relative PPP hypothesis[17]

$$\Delta\hat{R}_{ft} = a + b\Delta s + \epsilon$$

where H_0: $a = 0$ and $b = 1$ under the null hypothesis that relative PPP holds and the mean real rate differential is zero (see equation [7]), is tested against the specific alternative $H_a : \text{b} = 0$, i.e., spot rate changes are too volatile to be explained by inflation differentials.[18]

Before carrying out the PPP regression, we first test for unit roots in the foreign exchange rate changes and the extracted nominal risk-free rate differentials. Both the Augmented Dickey-Fuller and the Phillips-Perron tests reject the null of unit roots.

Table 7 reports relative PPP regression results for three different horizons: the monthly

[17] For comparison, we also used the CPI inflation differential as the dependent variable in the relative PPP regression. In this case, the intercepts are all significantly different from zero while slope coefficients are all close to zero and either not statistically significant or have the wrong sign. The adjusted R^2 are close to zero or negative. The detailed results are omitted from the paper for brevity and are available upon request.

[18] Note that the relative PPP hypothesis is tested against a specific alternative. The presence of a foreign exchange premium, as documented in some empirical studies, is also inconsistent with relative PPP in the sense that under this alternative $b \neq 1$. However, we do not know what value of b is implied by the presence of a foreign exchange premium, so a failure to reject the null of $b = 1$ may not shed much light on this question.

TABLE 5—TIME SERIES REGRESSIONS OF EXCESS INDUSTRIAL RETURNS ON THE THREE FAMA-FRENCH FACTORS (JAPAN)

Industry	Constant		$R_m - Tbill$		R_{SMB}		R_{HML}		Adj. R^2
Air transport	−0.200	(0.46)	0.843	(0.11)	0.439	(0.18)	−0.002	(0.20)	0.342
Banks	−0.003	(0.39)	1.173	(0.10)	0.207	(0.12)	−0.054	(0.14)	0.623
Chemical	−0.174	(0.21)	0.953	(0.05)	0.330	(0.09)	−0.025	(0.10)	0.757
Communication	0.833	(0.56)	1.113	(0.10)	−0.102	(0.13)	−0.514	(0.20)	0.450
Construction	−0.475	(0.35)	0.979	(0.09)	0.674	(0.17)	0.101	(0.13)	0.617
Electronic equipment	0.179	(0.39)	0.876	(0.09)	−0.475	(0.13)	0.323	(0.20)	0.462
Utilities	0.025	(0.45)	0.806	(0.09)	−0.291	(0.11)	−0.012	(0.20)	0.360
Fisheries	−0.465	(0.28)	0.854	(0.07)	0.901	(0.12)	−0.040	(0.14)	0.609
Foods	−0.167	(0.26)	0.791	(0.06)	0.337	(0.10)	−0.055	(0.10)	0.668
Glass & ceramics	−0.370	(0.26)	0.947	(0.06)	0.271	(0.08)	−0.047	(0.11)	0.717
Insurance	−0.049	(0.35)	1.132	(0.08)	−0.011	(0.11)	−0.028	(0.14)	0.639
Steel	−0.552	(0.41)	1.101	(0.05)	0.335	(0.12)	0.276	(0.18)	0.606
Land transport	0.054	(0.37)	0.887	(0.09)	0.220	(0.12)	−0.121	(0.17)	0.500
Machinery	−0.310	(0.26)	0.941	(0.07)	0.421	(0.10)	0.036	(0.11)	0.721
Marine transport	−0.603	(0.41)	1.143	(0.11)	0.784	(0.15)	−0.003	(0.18)	0.565
Metal products	−0.161	(0.30)	0.819	(0.07)	0.754	(0.12)	0.080	(0.14)	0.596
Mining	−0.628	(0.42)	1.004	(0.07)	0.894	(0.12)	−0.171	(0.17)	0.519
Non-ferrous metals	−0.450	(0.24)	1.081	(0.06)	0.241	(0.09)	−0.011	(0.11)	0.727
Oil & coal	−0.638	(0.33)	0.941	(0.09)	0.495	(0.17)	0.138	(0.21)	0.521
Other financials	−0.168	(0.37)	0.944	(0.06)	0.350	(0.16)	−0.163	(0.13)	0.598
Other products	0.046	(0.28)	0.716	(0.06)	0.097	(0.14)	0.030	(0.11)	0.514
Pharmaceutical	0.083	(0.31)	0.696	(0.07)	0.031	(0.11)	0.013	(0.14)	0.444
Precision instruments	0.094	(0.38)	0.777	(0.10)	−0.074	(0.10)	0.144	(0.17)	0.395
Paper	−0.276	(0.28)	0.789	(0.07)	0.615	(0.10)	0.019	(0.15)	0.564
Real estate	−0.313	(0.41)	1.248	(0.11)	0.020	(0.14)	0.189	(0.19)	0.561
Retail	0.409	(0.30)	0.800	(0.06)	0.206	(0.12)	−0.117	(0.11)	0.543
Rubber	0.103	(0.38)	0.916	(0.06)	0.224	(0.18)	0.119	(0.15)	0.555
Securities	−0.017	(0.38)	1.669	(0.15)	−0.140	(0.14)	0.239	(0.13)	0.732
Service	0.676	(0.38)	0.859	(0.07)	0.284	(0.15)	−0.339	(0.15)	0.565
Textiles	−0.462	(0.21)	0.900	(0.04)	0.544	(0.10)	0.046	(0.11)	0.738
Transport equipment	0.024	(0.26)	0.852	(0.05)	−0.472	(0.10)	0.455	(0.12)	0.632
Warehouse	−0.296	(0.35)	1.002	(0.07)	0.724	(0.13)	0.109	(0.17)	0.625
Wholesale	0.003	(0.29)	1.102	(0.06)	0.185	(0.13)	0.007	(0.14)	0.752

Notes: This table reports the regression results of excess industry portfolio returns on the three Fama-French factors for Japan with sample period from May 1983 to December 1999. The Newey-West adjusted standard errors are reported in parentheses to the right of the coefficient estimates.

TABLE 6—SUMMARY STATISTICS FOR EXTRACTED RISK-FREE RATE DIFFERENTIALS, CPI INFLATION DIFFERENTIALS, AND FOREIGN EXCHANGE RATE CHANGES

		CPI inflation differentials				Extracted risk-free rate differentials			Spot rate changes	
Currency pair	Frequency	Mean	Standard deviation	Correlation with extracted counterpart	Correlation with spot rate changes	Mean	Standard deviation	Correlation with spot rate changes	Mean	Standard deviation
Japanese ¥-US $	Monthly	−0.167	0.499	0.100	−0.065	−0.186	10.437	0.176	−0.424	3.520
	Bimonthly	−0.334	0.822	0.139	−0.170	−0.371	15.199	0.269	−0.847	5.364
	Quarterly	−0.500	0.681	0.268	−0.090	−0.697	20.925	0.275	−1.290	6.775
UK £-US $	Monthly	0.074	0.482	0.100	−0.029	0.275	8.874	0.100	−0.066	3.142
	Bimonthly	0.149	0.701	0.116	−0.052	0.550	12.813	0.263	−0.133	4.983
	Quarterly	0.223	0.797	0.322	−0.075	0.825	17.084	0.201	−0.199	5.454
German DM-US $	Monthly	−0.073	0.296	0.160	0.127	0.005	8.584	0.234	0.148	3.095
	Bimonthly	−0.146	0.430	0.064	0.170	0.011	11.720	0.419	0.296	4.648
	Quarterly	−0.218	0.534	0.189	0.210	0.016	15.688	0.501	0.443	6.031

Notes: This table reports the mean and standard deviation of extracted risk-free rate differentials from the Fama-French Three-Factor Model, official (CPI) inflation differentials, and foreign exchange rate changes. The inflation differential is the difference between a foreign and the U.S. extracted risk-free rate rates. The foreign exchange rate is measured as foreign currency per U.S. dollar. The CPI inflation is calculated from the CPI index. The correlations between the variables are also reported.

TABLE 7—PPP REGRESSION RESULTS USING EXTRACTED RISK-FREE RATE FROM THE FAMA-FRENCH THREE-FACTOR MODEL

			OLS results			Bootstrap results		$H_0: a = 0$ and $b = 1$	
Currency pair	Frequency	n	a	b	Adj. R^2	a	b	$F(2, n-2)$	p-value
Japanese ¥-US $	Monthly	200	0.036	0.522	0.026	0.028	0.535	2.706	0.069
			(0.639)	(0.250)*		(0.700)	(0.220)*		
	Bimonthly	100	0.276	0.763	0.063	0.281	0.769	0.421	0.658
			(1.021)	(0.337)*		(1.473)	(0.294)*		
	Quarterly	66	0.398	0.849	0.061	0.411	0.882	0.111	0.895
			(1.404)	(0.395)*		(2.343)	(0.382)*		
UK £-US $	Monthly	168	0.294	0.283	0.004	0.292	0.275	5.534	0.005
			(0.732)	(0.189)		(0.669)	(0.219)		
	Bimonthly	84	0.640	0.677	0.058	0.632	0.672	0.820	0.444
			(1.436)	(0.150)*		(1.372)	(0.220)*		
	Quarterly	56	0.950	0.631	0.023	0.892	0.578	0.494	0.613
			(2.254)	(0.352)**		(2.186)	(0.412)		
German DM-US $	Monthly	144	−0.091	0.648	0.048	−0.087	0.651	1.230	0.295
			(0.667)	(0.227)*		(0.686)	(0.212)*		
	Bimonthly	72	−0.302	1.056	0.164	−0.318	1.073	0.047	0.954
			(1.209)	(0.276)*		(1.222)	(0.298)*		
	Quarterly	48	−0.562	1.303	0.235	−0.631	1.302	0.440	0.647
			(1.587)	(0.380)*		(1.979)	(0.320)*		

Notes: This table reports the results of testing the relative PPP hypothesis by regressing the estimated inflation differential on the change in the nominal exchange rate, $R^*_{ft} - R_{ft} = a + b\Delta s_t + \epsilon_t$. The change in the foreign exchange rate is measured as yen or British pounds or Deutsche marks per U.S. dollar. Newey-West adjusted standard errors are reported in parentheses. n stands for the number of observations. The bootstrap point estimates and the standard errors are reported as well. The p value refers to the p-value of the joint F-test for the null hypothesis: intercept = 0 and slope = 1. * significantly different from zero at 5 percent; ** significantly different from zero at 10 percent.

horizon, the bimonthly horizon, and the quarterly horizon. Both the bimonthly and the quarterly horizons use non-overlapping observations. In order to adjust for the impact of heteroskedasticity and serial correlation, Newey-West adjusted standard errors are reported in the table.

At the monthly frequency, the coefficients for the contemporaneous exchange rate changes are significantly different from zero at the 5-percent level but not significantly different from one for the yen-dollar and the mark-dollar pair.

At the bimonthly and quarterly frequencies, the results are broadly consistent with the PPP hypothesis for all three pairs of currencies. The intercepts in all regressions are not significantly different from zero. The slopes are all above 0.6 and significantly different from zero, but not significantly different from one at the bimonthly frequency. At the quarterly frequency, the U.K.-U.S. pair has a slope estimate significantly different from zero at 10 percent but not from one at the same significance level. The adjusted R^2s, though not high, are not very different from our simulation results (omitted from the paper for brevity), especially for the mark-dollar pair. The small R^2 is consistent with our observation that there are large measurement errors contained in our nominal rate estimates.

For each individual country pair, an F test is carried out to test the joint hypothesis: $a = 0$ and $b = 1$.[19] Consistent with the individual t-tests, the test fails to reject the null that relative PPP holds in seven out of nine cases. It is strongly rejected with a p-value well below 1 percent for the pound-dollar pair at the monthly frequency and also rejected at the 10-percent significance level with a p-value around 7 percent for the yen-dollar pair at the monthly frequency, but the other seven tests have quite large p-values and fail to reject the null at a comfortable margin.

In contrast to the existing empirical results in the literature on testing the short-run PPP hypothesis, our point estimates are quite close to one in magnitude and generally significantly different from zero but not from one. Since the strongest evidence in favor of the relative PPP

[19] A joint test across all three country pairs is performed and reported in Section III C.

comes from regressions using bimonthly and quarterly data with relatively fewer observations, we also carry out a bootstrap simulation and report the point estimates and the bootstrapped standard errors in the same table.[20] Evident in both the point estimates and the standard errors, there is virtually no difference between the bootstrap results and the OLS estimates and the Newey-West standard errors: in the pound-dollar pair, the bootstrap standard errors are slightly larger and thus push the slope estimate from marginal significance to insignificance at the quarterly frequency, while the standard errors in the other two country pairs are smaller and thus strengthen our earlier results.

Finally, we carry out the PPP test for other cross-currency pairs, whose results are not reported for brevity. In summary, the PPP regression estimates of the slope coefficient are similar to those found in the results reported for currency pairs involving the U.S. dollar, but they are not statistically significant. It is worth mentioning here that the efficacy of these PPP tests depends critically on how well the Fama-French three-factor model describes the real stock returns and spans the real impact of inflation. We believe, however, that real factors left over in our estimates are more likely to *bias against* the PPP hypothesis, because there is no a priori theory or economic intuition that the spot rate changes move one-to-one with real factor *differentials* across different countries. To the extent that real factors may affect different economies at different times and with different significance, the spot rate changes are unlikely to move with the real factor differentials in a one-to-one relation across almost all country pairs that we have examined. The Fama-French three-factor model seems to fit the U.S. data quite well and has been subjected to intensive study; other countries, however, have not been scrutinized as thoroughly. Therefore, the stronger results for currency pairs involving the U.S. dollar are likely due to the better fit of the Fama-French three-factor model for U.S. portfolio returns.

B. *Results under the Four-Factor Model*

Some scholars have argued that the three-factor model of Fama and French does not adequately capture the time series variation in stock returns and a fourth real factor, momentum, explains a significant portion of stock returns.[21] We have the momentum factor only for the United States, so we repeated the analysis by fitting a four-factor model to U.S. industry portfolios. For brevity, in Table 8 we omit the time-series regression results and report only the PPP regression results.

The results support the PPP hypothesis even more strongly. In particular, the slope coefficients become closer to one with smaller standard errors, and the overall fit as measured by R^2s also improves at all frequencies.

The results remain virtually unchanged for the mark-dollar pair for which the original results in Table 7 were already quite strong. The slope estimates for the yen-dollar and the pound-dollar pairs improve to, respectively, 0.87 and 0.74 at the bimonthly frequency, and to 0.90 and 0.76 at the quarterly frequency. Although the point estimate $\hat{b}$ for the yen-dollar pair at the monthly frequency remains at the midway (0.58) between zero and one, the joint F-test now has a p-value of around 14 percent, compared to 7 percent before, and fails to reject the null hypothesis at the 10-percent significance level. The point estimate $\hat{b}$ for the pound-dollar pair at the monthly frequency remains low at only 0.33 and remains significantly different from one, but it is now also different from zero at a 10-percent significance level. Similar to the earlier observation, the bootstrap results are virtually the same as those of the OLS regressions.

The above results confirm our conjecture that real factors left over in our estimates are more

[20] In the bootstrap procedure, we randomly draw data from the original sample with replacement and carry out the regression using the new sample. The procedure is repeated 1,000 times, and the mean and the standard deviation across these 1,000 estimates are reported as the bootstrap point estimate and the bootstrap standard error in the table. See Bradley Efron and Robert J. Tibshirani (1994) for an excellent description of the bootstrap procedure.

[21] Carhart (1997) uses a four-factor model to evaluate mutual fund performance and argues that the four-factor model noticeably reduces the average pricing errors relative to both the CAPM and the three-factor model.

TABLE 8—PPP REGRESSION RESULTS USING EXTRACTED RISK-FREE RATE WITH MOMENTUM AS AN ADDITIONAL U.S. FACTOR

			OLS results			Bootstrap results		$H_0 : a = 0$ and $b = 1$	
Currency pair	Frequency	n	a	b	Adj. R^2	a	b	$F(2, n - 2)$	p-value
Japanese ¥-US $	Monthly	200	0.059	0.577	0.031	0.036	0.589	2.005	0.137
			(0.678)	(0.259)*		(0.740)	(0.228)*		
	Bimonthly	100	0.364	0.867	0.074	0.356	0.881	0.152	0.859
			(1.146)	(0.355)*		(1.489)	(0.308)*		
	Quarterly	66	0.425	0.903	0.062	0.431	0.919	0.052	0.950
			(1.577)	(0.413)*		(2.450)	(0.425)*		
UK £-US $	Monthly	168	0.296	0.326	0.007	0.305	0.326	4.791	0.010
			(0.744)	(0.196)**		(0.689)	(0.225)		
	Bimonthly	84	0.648	0.742	0.067	0.581	0.752	0.541	0.584
			(1.458)	(0.146)*		(1.394)	(0.212)*		
	Quarterly	56	0.975	0.756	0.036	0.870	0.710	0.256	0.775
			(2.298)	(0.351)*		(2.323)	(0.430)**		
German DM-US $	Monthly	144	−0.091	0.650	0.048	−0.077	0.654	1.214	0.300
			(0.672)	(0.227)*		(0.690)	(0.212)*		
	Bimonthly	72	−0.302	1.055	0.163	−0.309	1.056	0.046	0.955
			(1.223)	(0.277)*		(1.216)	(0.312)*		
	Quarterly	48	−0.561	1.301	0.231	−0.565	1.311	0.428	0.650
			(1.587)	(0.380)*		(1.979)	(0.320)*		

Notes: This table reports the results of testing the relative PPP hypothesis by regressing the estimated inflation differential on the change in the nominal exchange rate, $R^*_{ft} - R_{ft} = a + b\Delta s_t + \epsilon_t$. The change in the foreign exchange rate is measured as yen or British pounds or Deutsche marks per U.S. dollar. Newey-West adjusted standard errors are reported in parentheses. n stands for the number of observations. The bootstrap point estimates and the standard errors are reported as well. The p value refers to the p-value of the joint F-test for the null hypothesis: intercept = 0 and slope = 1. * significantly different from zero at 5 percent; ** significantly different from zero at 10 percent.

likely to bias against the finding of PPP relation. And a better factor model, by more effectively purging real effects of the nominal equity returns, leads to cleaner and better (less noisy) estimates of the ex post nominal risk-free rate, and thus a sharper test of the PPP relation in the last step.

C. *Results from Pooled Regressions*

Although our individual currency-pair results taken together are broadly in favor of the relative PPP hypothesis, one important currency pair, namely the pound-dollar pair, consistently rejects the PPP at the monthly frequency. A joint test statistic across all three country pairs may help us interpret the results in a more unified way. To serve this purpose, we pool all three country pairs in a system of equations. The additional advantage of a pooled regression is its enhanced power by using information from other cross-country pairs.

First, an unconstrained seemingly unrelated regression (SUR) is computed:

$$\begin{bmatrix} R^*_{ft,\mathrm{JP}} - R_{ft,\mathrm{US}} \\ R^*_{ft,\mathrm{UK}} - R_{ft,\mathrm{US}} \\ R^*_{ft,\mathrm{Ger}} - R_{ft,\mathrm{US}} \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} + \begin{bmatrix} \Delta s_{t,\mathrm{JP}} & 0 & 0 \\ 0 & \Delta s_{t,\mathrm{UK}} & 0 \\ 0 & 0 & \Delta s_{t,\mathrm{Ger}} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} + \epsilon_t.$$

Table 9 reports the SUR feasible generalized least square (FGLS) regression results in panel A for the monthly, bimonthly, and quarterly horizon where the contemporaneous correlations between different country pairs are taken into account, not only in the standard errors but also in the point estimates.[22] Panel B of the

[22] John H. Cochrane (2001) argues that the GLS procedure is more efficient only if the error covariance matrix is correctly modeled and the regression is perfectly specified. Otherwise, the GLS is less robust than OLS, especially when the variance-covariance matrix has to be estimated. In particular, the FGLS can put unreasonable weight on a slightly mis-specified area of the model and give misleading

TABLE 9—SEEMINGLY UNRELATED PPP REGRESSIONS FOR THREE COUNTRY PAIRS

Panel A: SUR FGLS estimates

			Three-factor model			Momentum as a 4th factor		
Currency pair	Frequency	n	a	b	System R^2	a	b	System R^2
Japanese ¥-US $	Monthly	512	−0.032	0.362	0.016	−0.020	0.391	0.017
			(0.730)	(0.192)**		(0.753)	(0.196)*	
UK £-US $			0.349	0.152		0.359	0.174	
			(0.675)	(0.202)		(0.678)	(0.202)	
German DM-US $			−0.067	0.530		−0.086	0.516	
			(0.692)	(0.219)*		(0.687)	(0.217)*	
Japanese ¥-US $	Bimonthly	256	0.082	0.535	0.067	0.124	0.584	0.065
			(1.477)	(0.252)*		(1.556)	(0.258)*	
UK £-US $			0.749	0.491		0.775	0.501	
			(1.340)	(0.252)**		(1.371)	(0.251)**	
German DM-US $			−0.241	0.968		−0.251	0.943	
			(1.262)	(0.271)*		(1.264)	(0.270)*	
Japanese ¥-US $	Quarterly	170	0.295	0.661	0.085	0.327	0.686	0.077
			(2.503)	(0.341)**		(2.641)	(0.350)*	
UK £-US $			1.051	0.450		1.074	0.497	
			(2.236)	(0.391)		(2.309)	(0.395)	
German DM-US $			−0.433	1.147		−0.462	1.083	
			(1.978)	(0.326)*		(1.986)	(0.323)*	

Panel B: Hypothesis testing

				Three-factor model		Momentum as a 4th U.S. factor	
H_0	Frequency	Df_1	Df_2	$F(Df_1, Df_2)$	p-value	$F(Df_1, Df_2)$	p-value
$a_1 = a_2 = a_3 = 0$	Monthly	3	506	0.119	0.949	0.126	0.945
	Bimonthly	3	250	0.128	0.943	0.136	0.939
	Quarterly	3	164	0.094	0.964	0.101	0.959
$b_1 = b_2 = b_3 = 0$	Monthly	3	506	2.867	0.036	2.919	0.034
	Bimonthly	3	250	6.247	<0.001	6.129	<0.001
	Quarterly	3	164	5.220	0.002	4.779	0.003
$b_1 = b_2 = b_3$	Monthly	2	506	0.911	0.403	0.785	0.457
	Bimonthly	2	250	0.981	0.376	0.788	0.456
	Quarterly	2	164	1.032	0.359	0.742	0.478
$b_1 = b_2 = b_3 = 1$	Monthly	3	506	8.865	<0.001	8.276	<0.001
	Bimonthly	3	250	2.109	0.100	1.816	0.145
	Quarterly	3	164	0.943	0.422	0.730	0.535
$a_1 = a_2 = a_3 = 0$	Monthly	6	506	4.507	<0.001	4.217	<0.001
and $b_1 = b_2 =$	Bimonthly	6	250	1.126	0.347	0.982	0.438
$b_3 = 1$	Quarterly	6	164	0.527	0.787	0.423	0.863

Notes: This table reports the results of testing the relative PPP hypothesis by using the seemingly unrelated system equations (SUR) across all three country pairs. The estimated inflation differential is regressed on the change in the nominal exchange rate, $\begin{bmatrix} R^*_{ft,\mathrm{JP}} - R_{ft,\mathrm{US}} \\ R^*_{ft,\mathrm{UK}} - R_{ft,\mathrm{US}} \\ R^*_{ft,\mathrm{Ger}} - R_{ft,\mathrm{US}} \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} + \begin{bmatrix} \Delta s_{t,\mathrm{JP}} & 0 & 0 \\ 0 & \Delta s_{t,\mathrm{UK}} & 0 \\ 0 & 0 & \Delta s_{t,\mathrm{Ger}} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} + \epsilon_t$. The change in the foreign exchange rate is measured as yen or British pounds or Deutsch marks per U.S. dollar. Newey-West adjusted standard errors are reported in parentheses, and n stands for the number of observations for the system. * significantly different from zero at 5 percent; ** significantly different from zero at 10 percent.

point estimates. The OLS estimates under SUR are the same as those under individual country-pair regressions reported in Tables 7 and 8.

table reports the F-statistics and the corresponding p-values for joint tests across equations. There are five null hypothesis tests in the table: (1) the intercepts are zero for all equations; (2) the slope coefficients are zero for all equations;

(3) the slope coefficients are the same across all equations; (4) the slope coefficients are all equal to one; and (5) the intercepts are zero and the slopes are one for all country pairs.

Comparing the results in panel A of Table 9 with those in Tables 7 and 8, we note that the FGLS slope coefficients, although still significant for the yen-dollar (at the 10-percent level) and mark-dollar (at the 5-percent level) pairs at the monthly horizon, are uniformly smaller in magnitude and closer to zero than to one when systems of equations are used. The results for the bimonthly horizon, however, are much more encouraging. Although the FGLS point estimates for the three slopes are also smaller than those reported in Tables 7 and 8, they are mostly above 0.5 and statistically different from zero but not from one, except for the pound-dollar pair under the three-factor model. The result for the mark-dollar pair remains the strongest support for the relative PPP where the slope estimate is around 0.97 under the three-factor model and 0.94 under the four-factor model. The results for the quarterly horizon show some slight improvement over the bimonthly results and are also in favor of the relative PPP hypothesis as illustrated by the individual t-ratios for the slope estimates.

The joint F tests reported in panel B of Table 9 lend strong support to the relative PPP hypothesis at the bimonthly and quarterly horizons as well. The p-values for the hypothesis that intercepts are zero for all equations are above 0.9 at all three horizons, so that the null is accepted. The null hypothesis that the slopes are all zero is rejected at a 5-percent-or-better significance level. One can easily accept the null that the slopes are all equal across the three equations, and the test fails to reject the null that the slopes are all equal to one as well, except for the monthly horizon. Finally, the null that the intercepts are all zero and the slopes are all one is also accepted at both the bimonthly and the quarterly horizons.

Since the hypotheses that the intercept, a_i, and the slope, b_i, are the same across equations are not rejected, these constraints are imposed in the following regression:

$$\begin{bmatrix} R^*_{ft,\mathrm{JP}} - R_{ft,\mathrm{US}} \\ R^*_{ft,\mathrm{UK}} - R_{ft,\mathrm{US}} \\ R^*_{ft,\mathrm{Ger}} - R_{ft,\mathrm{US}} \end{bmatrix} = a + \begin{bmatrix} \Delta s_{t,\mathrm{JP}} \\ \Delta s_{t,\mathrm{UK}} \\ \Delta s_{t,\mathrm{Ger}} \end{bmatrix} b + \epsilon_t .$$

Imposition of the constraint provides a more precise estimate of the coefficients and more powerful tests of the hypothesis if the constraints are true.

The results are reported in Table 10, where panel A contains results using the Fama-French three-factor model and panel B contains results when a four-factor model is applied to the U.S. data. The constrained intercept estimate, $\hat{a}$, is not significant, and the constrained slope estimate, $\hat{b}$, is significantly different from zero at all three horizons. Although $\hat{b}$ at the monthly horizon remains significantly different from one, $\hat{b}$'s at bimonthly and quarterly frequencies are quite close to one in magnitude and statistically insignificantly different from one. For example, $\hat{b}$ is above 0.9 at quarterly horizon under both the three-factor and four-factor models. The joint F-test for $H_0: a = 0,\ b = 1$ is also reported. Consistent with the implications from the individual t-ratios, the null hypothesis is strongly rejected at the monthly horizon but is comfortably accepted at the bimonthly and quarterly frequency.

Although the Newey-West standard errors adjust for heteroskedasticity and serial autocorrelation, they may still be underestimated due to the possible contemporaneous covariation among the three country pairs. To the extent that the three country pairs are perfectly contemporaneously correlated, the standard errors are underestimated by $\sqrt{3}$. A bootstrap simulation is carried out to obtain the contemporaneous-covariation-adjusted standard errors.

The system of equations has a total number of $n = 512$ observations across the three currency pairs. Instead of randomly drawing 512 observations from the original sample with replacement, only 144 observations, which is the sample size of the mark-dollar pair, are drawn from the original sample with replacement. The estimation is then carried out using the new sample. This procedure is repeated 2,000 times, and the average and the standard deviation across these 2,000 estimates are reported as the bootstrap estimate and standard errors in the table. The reported standard errors thus represent upper bounds. Note that the standard errors are substantially larger than the Newey-West ones, but the point estimates are very close to the OLS estimates. Despite the much larger standard errors, the null hypothesis $b = 1$ is still not rejected at the bimonthly and quarterly

TABLE 10—CONSTRAINED SYSTEM OF EQUATIONS PPP REGRESSION FOR THREE COUNTRY PAIRS

			OLS results			Bootstrap results		$H_0: a = 0$ and $b = 1$	
Currency pair	Frequency	n	a	b	Adj. R^2	a	b	$F(2, n-2)$	p-value
			Panel A: Three-factor model						
System of equations	Monthly	512	0.089	0.482	0.026	0.034	0.497	8.644	0.0002
			(0.389)	(0.141)*		(0.769)	(0.246)*		
	Bimonthly	256	0.272	0.803	0.087	0.261	0.820	0.847	0.430
			(0.753)	(0.175)*		(1.499)	(0.321)*		
	Quarterly	170	0.546	0.903	0.109	0.531	0.907	0.244	0.784
			(1.139)	(0.222)*		(2.291)	(0.396)*		
			Panel B: Four-factor model						
System of equations	Monthly	512	0.095	0.520	0.030	0.114	0.525	7.171	0.0008
			(0.400)	(0.144)*		(0.805)	(0.254)*		
	Bimonthly	256	0.291	0.869	0.095	0.276	0.878	0.394	0.675
			(0.781)	(0.182)*		(1.551)	(0.331)*		
	Quarterly	170	0.527	0.988	0.121	0.557	0.981	0.094	0.910
			(1.189)	(0.226)*		(2.323)	(0.402)*		

Notes: This table reports the results of testing the relative PPP hypothesis of the constrained system equations across all three country pairs. The estimated inflation differential is regressed on the change in the nominal exchange rate, $\begin{bmatrix} R^*_{ft,\mathrm{JP}} - R_{ft,\mathrm{US}} \\ R^*_{ft,\mathrm{UK}} - R_{ft,\mathrm{US}} \\ R^*_{ft,\mathrm{Ger}} - R_{ft,\mathrm{US}} \end{bmatrix} = a + \begin{bmatrix} \Delta s_{t,\mathrm{JP}} \\ \Delta s_{t,\mathrm{UK}} \\ \Delta s_{t,\mathrm{Ger}} \end{bmatrix} b + \epsilon_t$. The change in the foreign exchange rate is measured as yen or British pounds or Deutsche marks per U.S. dollar. Newey-West adjusted standard errors are reported in parentheses. The bootstrap point estimates and the standard errors are reported as well, and n stands for the number of observations for the system. For the bootstrap procedure, to guard against the possibility of underestimation of standard errors because of potential cross correlation among the three country pairs, only $n_3 < n/3$, the total number of observations for the Germany-U.S. pair, observations are drawn from the original sample with replacement; thus reported standard errors represent upper bounds. The p value refers to the p-value of the joint F-test for the null hypothesis: intercept = 0 and slope = 1. * significantly different from zero at 5 percent; ** significantly different from zero at 10 percent.

horizons while the null $b = 0$ is rejected at the 5-percent significance level for all three horizons.

D. *Robustness Checks*

Are the Results Driven by Real Effects of Inflation?—In our model (2), there is an implicit assumption that either the inflation does not have real effects or its real effects are spanned by the Fama-French three factors. Boudoukh et al. (1994) argue that expected inflation may be correlated with future production and thus affect real stock returns. They find some interesting inflation beta patterns across industries in a regression without any other factors, but none of their beta estimates is statistically significant. We also examine their conjecture by augmenting our factor models with an additional expected inflation factor. The results are omitted from the paper for brevity. In general, we find that most of the expected inflation betas are statistically insignificant, and the PPP test results improve marginally.

Are the Results Driven by a World Factor?—We did not include a world factor in the asset pricing model for the reasons given in Section IA. This means that the factor-generating model could possibly be mis-specified in the sense that the residuals η_{it} from the time-series regressions in (3) or our estimates of R_{ft} contain a missing world factor. The same world factor, expressed in units of foreign currencies, could also be contained in the corresponding estimates of R^*_{ft}. This suggests the possibility that our PPP tests might merely be detecting a much weaker phenomenon, viz., the law of one price holds for the world factor. To check on this possibility, we performed the following procedure.

The Morgan Stanley Composite Index (MSCI),[23] is a widely-used proxy for a world factor. We

[23] The data are provided by the MSCI Web site http://www.msci.com/

regressed the residuals from the time series industry returns and our estimates of R_{ft} on MSCI returns. The results (not reported here for brevity) reveal that neither the residuals from time series regressions nor the estimates of R_{ft} for any country are related to the MSCI returns.

Are the Results Driven by Other Missing Factors?—Even though neither the world factor nor the real effect of inflation is behind the supportive evidence on the short-run PPP relation, it is still possible that our extracted series represents a factor other than inflation and the differentials of this mysterious factor across all three country pairs happen to move one-to-one with the corresponding foreign exchange rate changes.

Since $\hat{R}_{ft}$ is estimated as the intercept of the cross-sectional regression in the Fama-MacBeth approach, $\hat{R}_{ft}$ moves one-to-one with all industry portfolio returns by construction. In our view, the most natural interpretation of $\hat{R}_{ft}$ in the Fama-MacBeth regression is the ex post nominal risk-free rate, which is then used as a surrogate for pure price inflation. If $\hat{R}_{ft}$ measures a non-inflation factor, then it must be systematic and influence all industries equally. If such a factor existed, then it should be subsumed in the market return factor. To see this formally, imagine that a factor θ_t affects returns of all industries equally and the return-generating process is described by:

$$R_{it} = \theta_t + R_{ft} + \beta_i(R_{Mt} - R_{ft}) + \epsilon_{it}.$$

Take a weighted sum, where weights are market value weights, across all i to obtain:

$$\sum_i w_i R_{it} = \theta_t + R_{ft} + (R_{Mt} - R_{ft}) \sum_i w_i \beta_i + \sum_i w_i \epsilon_{it}.$$

By definition $\Sigma_i w_i R_{it} = R_{Mt}$ and $\Sigma_i w_i \beta_i = 1$. For a large number of stocks or industry portfolios, $\Sigma_i w_i \epsilon_{it} \simeq 0$ which implies that

$$R_{Mt} \simeq \theta_t + R_{ft} + (R_{Mt} - R_{ft})$$

or that $\theta_t \simeq 0$. Therefore, it is unlikely that $\hat{R}_{ft}$ captures a systematic risk factor after controlling for the market factor in the regression.[24]

Since our estimates of the ex post nominal risk-free rates include the real risk-free rates, one might also argue that it is the real interest rate differential, rather than the pure price inflation differential, that is correlated with foreign exchange rate changes. Even if one could conceive of a model in which such a link can be established, it will be difficult to argue on *a priori* theoretical grounds that the volatility of *real* interest rate differentials could be of the same order of magnitude as the empirical estimates of the volatility of the nominal exchange rate changes.

Finally, it is unlikely that the extracted factor, if it measures something other than inflation and real risk-free rate, satisfies relative PPP so well. We further examine this issue in the next subsection by providing evidence that $\hat{R}_{ft}$ and the official CPI inflation measure have a long-run cointegration relation (Table 11) even though their short-run correlation is weak (Table 6).

How Well Does $\hat{R}_{ft}$ *Track the Official Inflation Measure?*—We examine what relation, if any, exists between our extracted nominal risk-free rates and official inflation estimates. A direct comparison is difficult for five reasons. One, our extracted variables $\hat{R}_{ft}$ contain both pure price inflation and real interest rates. That's why the sample means of our extracted variables are higher than those of CPI inflation measures, π_{CPI}, as reported in Table I. Two, our $\hat{R}_{ft} = R_{ft} + \epsilon_t$'s are estimated with considerable noise, as illustrated by their large sample volatilities. Three, the CPI is constructed from commodity prices sampled throughout the month, and it provides a closer measure of inflation from mid-month to mid-month while the extracted risk-free rate is the end-of-month measure, leading to some mismatching of time period. Four, inflation measures using CPI data would include not only pure price inflation but also effects of relative price changes. This

[24] Because of this argument, it is unlikely that $\hat{R}_{ft}$ contains a foreign exchange rate factor. The existing empirical evidence of whether foreign exchange risk affects expected returns is mixed. The foreign exchange rate is found to be a significant factor only for a subset of firms and its beta is rarely close to one. G. Andrew Karolyi and Rene M. Stulz (2003) provide a survey of the empirical findings.

TABLE 11—COINTEGRATION RELATION BETWEEN THE CPI AND THE CONSTRUCTED PRICE INDEX

	Cointegration regression			Johansen test
Currency	Intercept	$\hat{c}_1$	LL	Null: no cointegration
Panel A: Price indices				
GM	−1.842	1.672	778.0	Reject
	[1.75]	[2.07]		
JP	0.368	0.318	986.3	Reject
	[0.43]	[0.43]		
UK	−10.972	4.806	639.7	Reject
	[2.96]	[2.24]		
US	−1.082	0.989	1088.2	Reject
	[2.88]	[3.94]		
Panel B: Logged price indices				
GM	−1.434	2.747	824.1	Reject
	[2.45]	[4.40]		
JP	−0.395	0.686	1025.4	Reject
	[3.06]	[0.80]		
UK	−7.745	9.196	848.8	Reject
	[2.96]	[2.24]		
US	−0.899	1.658	1220.7	Reject
	[7.19]	[7.24]		

Notes: This table reports the Johansen test of cointegration between the price index, $P_{R_f,t}$, which is constructed from the extracted series $\hat{R}_{ft}$, and the official price index CPI. The cointegration regression results are also reported with the asymptotic *t*-ratios in the brackets. Panel A reports the analysis using the price indices themselves while Panel B reports the results using logged price indices.

reason is one of the motivations for constructing a pure inflation measure. Finally, as we argued earlier, our extracted price inflation measure is responsive to news and changes in expectations about future prices, whereas official indices may be sticky.

Nevertheless, we expect some association between $\hat{R}_{ft}$ and π_{CPI} in the long run when the temporary swings in short-run inflation cancel one another out and the smooth official series becomes a better measure of realized long-run inflation. To the extent that they both measure inflation well over the long run, the price level implied by the two series should be cointegrated.[25]

To examine this hypothesis, we first create a hypothetical price index for $\hat{R}_{ft}$, $P_{R_f,t}$, by setting the initial value to one ($P_{R_f,0} = 1$):

$$P_{R_f,t} = P_{R_f,t-1}(1 + \hat{R}_{ft}).$$

The CPI is similarly constructed:

$$P_{\text{CPI},t} = P_{\text{CPI},t-1}(1 + \pi_{\text{CPI}})$$

which is equivalent to resetting the official CPI to one at the beginning of the sample.

The Augmented Dickey-Fuller (ADF) test cannot reject the null of unit root for all P_{CPI} series across the four countries and for P_{R_f} in Germany, the United Kingdom, and the United States at the 5-percent level, but it does reject the null of unit root for the Japanese P_{R_f} (p-value $= 0.045$). On the other hand, the ADF fails to reject the null of unit root in all logged price indices. In other words, all price level series appear to be non-stationary.

Table 11 reports the result of the Johansen cointegration test and the Johansen Maximum Likelihood estimate of the cointegrating vector, with the coefficient of P_{R_f} normalized to unity. The Johansen test rejects the null of no cointegration between P_{R_f} and the CPI across all four countries, indicating that there is a long-run cointegration relation between P_{R_f} and CPI. The cointegration regression yields highly significant cointegrating parameters in three currencies (GM, UK and US). Although the cointegrating coefficient need not equal one for reasons mentioned above, it is actually not significantly different from one for all four countries.

IV. Concluding Remarks

Our paper makes two contributions. First, we provide a novel application of the Fama-MacBeth approach to extract the ex post nominal risk-free return from stock returns, which is then used as a surrogate for *realized* pure price inflation. This is done by exploiting the fact that the intercept in the second-stage Fama-MacBeth procedure measures the nominal risk-free rate. Instead of focusing on the sample mean of this series as is often done in the asset pricing literature, we take advantage of its time series behavior and use it as a surrogate for the unobservable inflation series.

Second, we provide compelling evidence that relative purchasing power parity holds quite well in the short run when inflation is extracted

[25] We are grateful to the editor, Ben Bernanke, for suggesting that we test for cointegration between our extracted series and the official CPI series.

from stock prices. Our results complement the current consensus that PPP holds in the long run. This is in sharp contrast to the poor performance of short-term PPP documented in the extensive literature surveyed in Rogoff (1996) where inflation measures were based on official series such as the CPI.

The support for PPP using our extracted inflation measures suggests that the price level affecting financial markets is much more volatile than official price indices computed from the prices of goods and services. It is natural to ponder why monetary shocks might affect the financial market and the goods and services market in such different ways. One possibility is that financial asset prices are affected by "news" and "perceptions" which may be temporary in nature and may not materialize in the future, while prices of consumer goods change gradually and infrequently only with clear and compelling evidence of a long-run permanent event. In addition, prices in financial markets are usually fully flexible and determined by continuous trading of investors, whereas prices of consumer goods are often posted before the transaction date; it can be costly to change a posted price despite the small "menu cost," as argued in Mankiw (1985).

Even though the short-run correlation between our extracted series and official CPI inflation is weak, we find strong evidence that our extracted price index and the CPI are cointegrated.

Our series reflects high volatility of inflation in the short run, which is not well captured by official CPI inflation. If the inflation reflected in financial markets is much more volatile than official inflation, there is no need to rely on some form of exchange rate "overshooting" (Dornbusch, 1976) to explain exchange rate volatility, particularly given that implications of overshooting do not seem to be supported by empirical evidence (Rogoff, 2002).

REFERENCES

Adler, Michael and Lehmann, Bruce. "Deviations from Purchasing Power Parity in the Long Run." *Journal of Finance*, 1983, *38*(5), pp. 1471–87.

Ang, Andrew and Bekaert, Geert. "Stock Return Predictability: Is It There?" National Bureau of Economic Research, Inc., NBER Working Papers: No. 8207, 2001.

Barr, David G. and Campbell, John Y. "Inflation, Real Interest Rates, and the Bond Market: A Study of UK Nominal and Index-Linked Government Bond Prices." *Journal of Monetary Economics*, 1997, *39*(3), pp. 361–83.

Baxter, Marianne. "Real Exchange Rates and Real Interest Differentials: Have We Missed the Business-Cycle Relationship?" *Journal of Monetary Economics*, 1994, *33*(1), pp. 5–37.

Bleaney, Michael and Laxton, Douglas. "Real Interest Rates and Real Exchange Rates: Evidence from Indexed Bonds." *Manchester School*, 2003, *71*(1), pp. 65–77.

Boudoukh, Jacob; Richardson, Matthew and Whitelaw, Robert F. "Industry Returns and the Fisher Effect." *Journal of Finance*, 1994, *49*(5), pp. 1595–1615.

Carhart, Mark M. "On Persistence in Mutual Fund Performance." *Journal of Finance*, 1997, *52*(1), pp. 57–82.

Cochrane, John H. *Asset pricing.* Princeton: Princeton University Press, 2001.

Delgado, Francisco A. "Hysteresis, Menu Costs, and Pricing with Random Exchange Rates." *Journal of Monetary Economics*, 1991, *28*(3), pp. 461–84.

Dixit, Avinash K. "Entry and Exit Decisions under Uncertainty." *Journal of Political Economy*, 1989a, *97*(3), pp. 620–38.

Dixit, Avinash K. "Hysteresis, Import Penetration, and Exchange Rate Pass-Through." *Quarterly Journal of Economics*, 1989b, *104*(2), pp. 205–28.

Dornbusch, Rudiger. "Expectations and Exchange Rate Dynamics." *Journal of Political Economy*, 1976, *84*(6), pp. 1161–76.

Edison, Hali J. and Pauls, B. Dianne. "A Re-Assessment of the Relationship between Real Exchange Rates and Real Interest Rates: 1974–1990." *Journal of Monetary Economics*, 1993, *31*(2), pp. 165–87.

Efron, Bradley and Tibshirani, Robert J. *An introduction to the bootstrap.* Monographs on Statistics and Applied Probability, Vol. 57. London: Chapman and Hall/CRC LLC, 1993.

Fama, Eugene F. and MacBeth, James D. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy*, 1973, *81*(3), pp. 607–36.

Fama, Eugene F. and French, Kenneth R. "Size and Book-to-Market Factors in Earnings and Returns." *Journal of Finance*, 1995, *50*(1), pp. 131–55.

Griffin, John M. "Are the Fama and French Factors Global or Country Specific?" *Review of Financial Studies*, 2002, *15*(3), pp. 783–803.

Karolyi, G. Andrew and Stulz, Rene M. "Are Financial Assets Priced Locally or Globally?" in George Constantinides, Milton Harris, and René Stulz, eds., *The handbook of the economics of finance 1B.* Amsterdam: Elsevier North-Holland, 2003.

Krugman, Paul. "Pricing to Market when the Exchange Rate Changes." National Bureau of Economic Research, Inc., NBER Working Papers: No. 1926, 1986.

Mankiw N. Gregory. "Small Menu Costs and Large Business Cycles: A Macroeconomic Model of Monopoly." *Quarterly Journal of Economics*, 1985, *100*(2), pp. 529–38.

Mussa, Michael. "A Model of Exchange Rate Dynamics." *Journal of Political Economy*, 1982, *90*(1), pp. 74–104.

Newey, Whitney K. and West, Kenneth D. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 1987, *55*(3), pp. 703–08.

Pagan, Adrian. "Econometric Issues in the Analysis of Regressions with Generated Regressors." *International Economic Review*, 1984, *25*(1), pp. 221–47.

Rogoff, Kenneth. "The Purchasing Power Parity Puzzle." *Journal of Economic Literature*, 1996, *34*(2), pp. 647–68.

Rogoff, Kenneth. "Perspectives on Exchange Rate Volatility," in Martin Feldstein, ed., *International capital flows.* Chicago: University of Chicago Press, 1999, pp. 441–53.

Rogoff, Kenneth. "Dornbusch's Overshooting Model after Twenty-Five Years," Second Annual Research Conference, Mundell-Fleming Lecture. Washington, DC: International Monetary Fund, 2002.

Rotemberg, Julio J. "Sticky Prices in the United States." *Journal of Political Economy*, 1982, *90*(6), pp. 1187–211.

Shanken, Jay. "On the Estimation of Beta-Pricing Models." *Review of Financial Studies*, 1992, *5*(1), pp. 1–33.

Stock, James H. and Watson, Mark W. "Forecasting Output and Inflation: The Role of Asset Prices." *Journal of Economic Literature*, 2003, *41*(3), pp. 788–829.

Zhang, Xiaoyan. "Specification Tests of International Asset Pricing Models." *Journal of International Money and Finance* (forthcoming).

This article has been cited by:

1. Hyein Shim, Hyeyoen Kim, Sunghyun Kim, Doojin Ryu. 2016. Testing the relative purchasing power parity hypothesis: the case of Korea. *Applied Economics* **48**:25, 2383-2395. [Crossref]
2. Shiu-Sheng Chen. 2010. Does extracting inflation from stock returns solve the purchasing power parity puzzle?. *Empirical Economics* . [Crossref]
3. G CHORTAREAS, G KAPETANIOS. 2009. Getting PPP right: Identifying mean-reverting real exchange rates in panels. *Journal of Banking & Finance* **33**:2, 390-404. [Crossref]