Online Appendix
International Portfolio Choice with Frictions: Evidence from Mutual Funds

September 2020

This Online Appendix has four sections. Section A presents the return differential regressions when applied to our EPFR portfolio sample. Section B discusses portfolio regressions using data aggregated over funds. Section C presents results with alternative specifications of portfolio regressions. Section D discusses robustness analysis.

A Predicting Cross-Country Equity Return Differentials

The return differential regressions in Section 3 of the paper are done for 73 countries at different horizons. In the portfolio regressions using EPFR data, the number of countries is reduced to 35. Moreover, the expected return differentials are discounted by the factor δ that we wet at 0.89. Table A1 shows the results of the pooled regressions for the discounted return differential when using the 35 countries of the EPFR sample.

B Aggregate Portfolio Regressions

We consider portfolio regressions that use portfolio shares aggregated across all the funds. We show that this continues to deliver a large and significant weight on the average benchmark portfolio and a statistically significant coefficient on the expected excess return. However, the magnitudes of the coefficients differ substantially from those based on fund-level portfolio regressions. A major reason is endogeneity, as explained in the text. But there are several other reasons for these differences.

To examine this, we aggregate EPFR fund-level data by investment country. This gives us a panel with 35 investment countries for EPFR US mutual fund equity investors. More precisely, the country-level portfolio share $z_{n,t}$ is constructed as:

$$z_{n,t} = \sum_{i \in I_{n,t}} f_{i,t} z_{i,n,t} \quad \text{(B.1)}$$

Table A1: Regressions Return Differential - Different Horizons - Discounted - EPFR Sample

	(1) $er_{n,t,t+1}$	(2) $er^{0.89}_{n,t,t+12}$	(3) $er^{0.89}_{n,t,t+24}$	(4) $er^{0.89}_{n,t,t+60}$
Momentum	0.0378*	0.164***	0.168***	0.174***
	(0.0194)	(0.0428)	(0.0430)	(0.0444)
Dividend-Price	0.00447	0.0485***	0.0638***	0.0798***
	(0.00280)	(0.00669)	(0.00729)	(0.00767)
Earning-Price	0.00438**	0.0212***	0.0289***	0.0285***
	(0.00219)	(0.00521)	(0.00543)	(0.00566)
Constant	-0.000453	-0.00338	-0.00209	0.00259
	(0.00149)	(0.00396)	(0.00436)	(0.00468)
Observations	14672	14287	13867	12612
R^2	0.004	0.038	0.057	0.074

Standard errors clustered by month in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Notes: Regressions with 35 countries over the interval 1970:01-2019:02. All regressions include a country fixed effect.

where $I_{n,t}$ is the set of funds i that invest in country n at time t and $f_{i,t}$ is asset under management (AUM) of fund i divided by AUM of all funds. The aggregate buy-and-hold portfolio $z_{n,t}^{bh}$ is computed as $z_{n,t-1}$ times the gross stock return in country n, divided by the aggregate portfolio return with country weights $z_{m,t-1}$.

We compute the expected return differential $E_t er_{n,t,t+24}^{\delta}$ as in (12) in the text, but replacing the fund weights $\bar{z}_{i,m,-n}$ with the aggregate EPFR weights $\bar{z}_{m,-n}$. In other words, we compare the return in country n to the weighted average of returns of all other countries, using aggregate EPFR weights in the other countries. We then estimate

$$z_{n,t} = a_0 + a_1 \frac{z_{n,t-1} + z_{n,t}^{bh}}{2} + a_2(z_{n,t-1} - z_{n,t}^{bh}) + a_3 E_t er_{n,t,t+24}^{0.89} + \varepsilon_{n,t} \tag{B.2}$$

The first column of Table B1 (which corresponds to column 4 of Table 3 in the text) shows the result of estimating (B.2). While all coefficients are significant, their magnitude is quite different from the fund-level results in Table 3 in the paper. The estimate of the return sensitivity a_3 is much smaller than the return sensitivity b_3 for individual funds. The weight on the average portfolio share is 0.998 is much closer to 1 than the coefficient based on fund-specific data. Finally, the lagged portfolio share is less important than the buy-and-hold portfolio for aggregate data, while the opposite was the case for fund-level data. These differences cannot be explained by endogeneity alone since they are already large when we compare with column 1 of Table 3, where we do not correct for endogeneity.

What could explain the difference in results? There are two sets of explanations, one related to fund heterogeneity and another unrelated to heterogeneity. Starting with the latter, consider the portfolio regression (13) in the text and assume that the coefficients b_1, b_2 and b_3 are the same for all funds. Multiplying both sides of (13) in the text by $f_{i,t}$ and taking the sum over all funds in $I_{n,t}$, we get

$$z_{n,t} = b_0 + b_1 \frac{\hat{z}_{n,t-1} + \hat{z}_{n,t}^{bh}}{2} + b_2 \left(\hat{z}_{n,t-1} - \hat{z}_{n,t-1}^{bh}\right) + b_3 f_t^n E_t \hat{er}_{n,t,t+24}^{\delta} \tag{B.3}$$

where $\hat{z}_{n,t-1} = \sum_{i \in I_{n,t}} f_{i,t} z_{i,n,t-1}$, $\hat{z}_{n,t}^{bh} = \sum_{i \in I_{n,t}} f_{i,t} z_{i,n,t}^{bh}$, $f_t^n = \sum_{i \in I_{n,t}} f_{i,t}$, and

$$E_t \hat{er}_{n,t,t+24}^{\delta} = E_t \sum_{s=1}^{24} \delta^{s-1} er_{n,t+s} - \sum_{m \neq n} \hat{z}_{m,-n} E_t \sum_{s=1}^{24} \delta^{s-1} er_{m,t+s}$$

where $\hat{z}_{m,-n}$ is the weighted average of $\bar{z}_{i,m,-n}$, with fund weights based on the AUM of funds that invest in country n at time t.

Table B1: PORTFOLIO REGRESSIONS, COUNTRY-LEVEL AGGREGATE

	(1)	(2)	(3)	(4)
$(z_{n,t-1} + z^{bh}_{n,t})/2$	0.998***			
	(0.002)			
$(z_{n,t-1} - z^{bh}_{n,t})$	-0.217***			
	(0.012)			
$E_t(er^{0.89}_{n,t,t+24} - \sum_m \bar{z}_{m,-n} er^{0.89}_{m,t,t+24})$	0.026**	0.070***		
	(0.012)	(0.015)		
$(\hat{z}_{n,t-1} + \hat{z}^{bh}_{n,t})/2$		0.992***	0.990***	0.989***
		(0.002)	(0.002)	(0.002)
$(\hat{z}_{n,t-1} - \hat{z}^{bh}_{n,t})$		-0.063	-0.058	-0.057
		(0.039)	(0.039)	(0.039)
$f^n_t E_t(er^{0.89}_{n,t,t+24} - \sum_m \bar{z}_{m,-n} er^{0.89}_{m,t,t+24})$			0.333***	
			(0.069)	
$f^n_t E_t(er^{0.89}_{n,t,t+24} - \sum_m \hat{z}_{m,-n} er^{0.89}_{m,t,t+24})$				0.350***
				(0.073)
Observations	5918	5918	5918	5918
R^2	0.999	0.999	0.999	0.999

Standard errors clustered by month in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Notes: Country-level regressions over the interval 2002:01-2016:07. All regressions include a country fixed effect. The constant is included but not shown.

Equation (B.3) properly aggregates the fund-level portfolio expressions. If the coefficients b_i were the same for all funds, their estimates based on (B.3) should be consistent with those based on the fund-level portfolio shares reported in Table 3 in the paper. Estimating (B.3) requires the use of various fund-specific pieces of information about AUM and portfolio shares. An econometrician who only has access to aggregate portfolio data would instead estimate (B.2).

There are three reasons why the estimates of a_i would differ from those of b_i. First, the lagged portfolio share $z_{n,t-1}$ and buy-and-hold portfolio share $z_{n,t}^{bh}$ in (B.2) are replaced by respectively $\hat{z}_{n,t-1}$ and $\hat{z}_{n,t}^{bh}$ in (B.3). As shown in column 2 of Table B1 making these replacements leads to a coefficient on the difference between the benchmark portfolio shares that becomes insignificant. It also triples the coefficient on the expected excess return.

Second, the expected excess return variable in (B.3) is multiplied by f_t^n, the fraction of all AUM accounted for by funds that invest in country n. Even if the portfolios of all funds that invest in country n have a responsiveness to the expected excess return of b_3, the responsiveness of the aggregate portfolio is only $b_3 f_t^n$ as these funds make up a limited share of f_t^n of aggregate AUM. On average f_t^n is 0.325. One can therefore expect the coefficient on the expected excess return variable to rise once we replace the expected excess return in (B.2) with the same excess return times f_t^n. This is indeed what happens in column 3 of Table B1, where the coefficient on the expected excess return variable is almost five times that in column 2.

Finally, the expected excess return variable in the aggregate portfolio regression (B.2) is different from that in (B.3). The difference relates to the reference countries that the return in country n needs to be compared to. The proper aggregation of fund-specific portfolio regressions in (B.3) accounts for the fact that for example funds investing in France have a different set of reference countries than funds investing in Peru. This is incorporated in the reference country portfolio shares $\hat{z}_{m,-n}$. By contrast, the reference portfolio shares $\bar{z}_{m,-n}$ that are used to compute the expected excess return in (B.2) are simply based on the aggregate portfolio shares $\bar{z}_m$ (rescaled by subtracting country n). However, we find in column 4 of Table B1 that when we replace the weights $\bar{z}_{m,-n}$ with $\hat{z}_{m,-n}$, the coefficients remain virtually unchanged.

Column 4 of Table B1 reports estimates of b_1, b_2 and b_3 that should be comparable to those of Table 3, column 1, for fund-level portfolio regressions. However,

$b_1 = 0.989$ is still higher than 0.928 in Table 3 and $b_3 = 0.350$ remains lower than 1.082. The remaining differences can be attributed to parameter heterogeneity across funds. If the fund-specific coefficients are $b_{1,i}$, $b_{2,i}$ and $b_{3,i}$, the estimates in Table 3 should be interpreted as averages across funds.

To see what the implications are for such parameter heterogeneity, consider the fund level expression, but with heterogeneous coefficients:

$$z_{i,n,t} = b_0 + b_{in} + b_{1,i}\frac{z_{i,n,t-1} + z_{i,n,t}^{bh}}{2} + b_{2,i}\left(z_{i,n,t-1} - z_{i,n,t}^{bh}\right) + b_{3,i}E_t er_{i,n,t,t+k}^{\delta} + \varepsilon_{i,n,t} \tag{B.4}$$

Aggregate these in the same way that lead to (B.3). First multiply by $f_{i,t}$ and then sum over all i in $I_{n,t}$:

$$\begin{aligned} z_{n,t} \quad = \quad & b_0 + \sum_{i\in I_{n,t}} b_{1,i}f_{i,t}\frac{z_{i,n,t-1} + z_{i,n,t}^{bh}}{2} + \sum_{i\in I_{n,t}} b_{2,i}f_{i,t}\left(z_{i,n,t-1} - z_{i,n,t}^{bh}\right) + \\ & \sum_{i\in I_{n,t}} b_{3,i}f_{i,t}E_t er_{i,n,t,t+k}^{\delta} + \varepsilon_{n,t} \end{aligned} \tag{B.5}$$

The three summation terms can be treated analogously. To illustrate, take the first one. Define $x_{i,n,t} = f_{i,t}\frac{z_{i,n,t-1}+z_{i,n,t}^{bh}}{2}$. We have

$$\sum_{i\in I_{n,t}} b_{1,i}x_{i,n,t} = \left(\frac{1}{I_{n,t}}\sum_{i\in I_{n,t}} b_{1,i}\right)\left(\sum_{i\in I_{n,t}} x_{i,n,t}\right) + I_{n,t}cov(b_{1,i}, x_{i,n,t}) \tag{B.6}$$

where the covariance is the cross sectional covariance across i in $I_{n,t}$ between $b_{1,i}$ and $x_{i,n,t}$. Note that $\sum_{i\in I_{n,t}} x_{i,n,t} = (\hat{z}_{n,t-1} + \hat{z}_{n,t}^{bh})/2$. The average of the coefficients $b_{1,i}$ is b_1, which we estimated in the fund-level regressions. We then have

$$\sum_{i\in I_{n,t}} b_{1,i}x_{i,n,t} = b_1\frac{\hat{z}_{n,t-1} + \hat{z}_{n,t}^{bh}}{2}\left(1 + cov\left(\frac{b_{1,i}}{b_1}, \frac{f_{i,t}(z_{i,n,t-1} + z_{i,n,t}^{bh})}{(\hat{z}_{n,t-1} + \hat{z}_{j,n,t}^{bh})/I_{n,t}}\right)\right) \tag{B.7}$$

Doing this for all terms in (B.5), we get

$$\begin{aligned} z_{n,t} \quad = \quad & b_0 + b_1\frac{\hat{z}_{n,t-1} + \hat{z}_{n,t}^{bh}}{2}\left(1 + cov\left(\frac{b_{1,i}}{b_1}, \frac{f_{i,t}(z_{i,n,t-1} + z_{i,n,t}^{bh})}{(\hat{z}_{n,t-1} + \hat{z}_{n,t}^{bh})/I_{n,t}}\right)\right) \\ & + b_2(\hat{z}_{n,t-1} - \hat{z}_{n,t}^{bh})\left(1 + cov\left(\frac{b_{2,i}}{b_2}, \frac{f_{i,t}(z_{i,n,t-1} - z_{i,n,t}^{bh})}{(\hat{z}_{n,t-1} - \hat{z}_{n,t}^{bh})/I_{n,t}}\right)\right) \\ & + b_3 f_t^n E_t\hat{er}_{n,t,t+24}^{\delta}\left(1 + cov\left(\frac{b_{3,i}}{b_3}, \frac{f_{i,t}E_t er_{i,n,t,t+k}^{\delta}}{f_t^n E_t\hat{er}_{n,t,t+24}^{\delta}/I_{n,t}}\right)\right) \end{aligned} \tag{B.8}$$

We can conclude that the coefficients b_1, b_2 and b_3 in the regression (B.3) reported in the last column of Table B1 are biased upwards (in terms of measuring the average of the coefficients across funds) to the extent that funds with a relatively large coefficient are also large in AUM (large f_{it}). We saw in Section 5.2 that larger funds have a larger coefficient b_1 and a smaller coefficient b_3. This would bias upward the estimate of b_1 in the last column of Table B1, and bias downward the estimate of b_3.

C Alternative Specifications of the Portfolio Regression

Table C1: Portfolio Regressions, Alternative Specifications

	(1)	(2)	(3)	
$z_{i,n,t-1}$	0.771***	0.916***		0.916***
	(0.069)	(0.005)		(0.005)
$z^{bh}_{i,n,t}$	0.145**		0.897***	
	(0.069)		(0.008)	
$E_t er^{0.89}_{i,n,t,t+24}$	2.324***	2.401***	2.506***	2.446***
	(0.291)	(0.287)	(0.385)	(0.287)
$val_{i,n,t}$		0.004*	-0.061***	
		(0.002)	(0.007)	
Observations	316732	316732	316732	316732
R^2	0.988	0.988	0.987	0.988

Standard errors clustered by month in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Notes: Regressions with 35 countries over the horizon 2002:01-2016:07. All regressions include a fund-country fixed effect and a country-month fixed effect. The constant is included in the regressions but not shown.

Columns (1) to (3) Table C1 show regressions of alternative ways of writing the portfolio equation, as shown in the Appendix of the paper. We focus on the benchmark regression of column 2 in Table 3. These alternative regressions use directly the benchmark portfolios instead of their average and difference. Column (1) corresponds to equation (A.8) in the paper. Columns (2) and (3) correspond to equations (A.6) and (A.9), respectively. The variable $val_{i,n,t}$ is the valuation effect defined as the difference between net country returns and net fund returns. Raddatz and Schmukler (2012) refer to is as *Relative returns* and express it in decimal points. Columns (2) and (3) are similar to the specifications of Table 5 and Table 6 in Raddatz and Schmukler (2012) but include expected excess returns and country-month fixed effects.

Column (4) abstracts from the buy-and-hold portfolio. This case corresponds to what we consider in Section 4.3, when we assume $z_{i,n,t}^{bh} = z_{i,n,t-1}$.

D Robustness Analysis

Table D1 tests the benchmark portfolio equation over different data samples. Columns (1) and (2) are the benchmark equations when we start the sample in January 2010 and in January 2012. Column (3) restricts the sample to funds continuously reporting from January 2002 to July 2016. Column (4) restricts the sample to funds reporting from January 2010 and reporting for at least 70 months. Column (5) restricts the sample to funds investing at least for 24 consecutive months. While the precise estimates vary across samples, in all cases there is a coefficient of 0.89 or higher on the average portfolio, either the same or a higher relative weight on the lagged portfolio than the buy-and-hold portfolio, and a strongly significant coefficient on the expected excess return variable.

Table D2 tests the benchmark equations with different predicted excess returns. Column (1) only uses differentials in earning-price ratio with momentum. Column (2) uses differentials in dividend-price ratio with momentum. Column (3) uses differentials in momentum, dividend-price and earning-price ratio up to three lags to predict excess returns. Column (4) adds the one-month interest rate differential to the three benchmark variables (momentum, dividend-price, and earning-price). The results are very similar with the three-month interest rates.[1] Finally, column (5) uses differentials in momentum, dividend-price and earning-price ratio non-recursively, i.e., using the whole sample to estimate return differentials so that these are not true forecasts. None of these significantly change conclusions. It

[1]The interest rates correspond to one-month and three-month annual Eurorates provided by Intercapital from Datastream. The data are midpoint of the offer and bid rates. We construct the monthly series by using the last trading day of the monthly data. Original data are expressed at annual rates in percent, and we construct the monthly series by dividing by 1200.

Table D1: Portfolio Regressions, Alternative Samples

	From Jan, 2010	From Jan, 2012	Full Reporting	Full Reporting From Jan, 2010	Report more than 24 months
	(1)	(2)	(3)	(4)	(5)
$(z_{i,n,t-1} + z^{bh}_{i,n,t})/2$	0.897***	0.887***	0.963***	0.931***	0.923***
	(0.006)	(0.008)	(0.007)	(0.007)	(0.005)
$(z_{i,n,t-1} - z^{bh}_{i,n,t})$	0.349***	0.181***	0.118**	0.206***	0.320***
	(0.059)	(0.027)	(0.053)	(0.030)	(0.070)
$E_t er^{0.89}_{i,n,t,t+24}$	2.896***	2.827***	1.358**	2.709***	2.115***
	(0.365)	(0.359)	(0.561)	(0.346)	(0.287)
Observations	229038	190856	18792	81140	275740
R^2	0.988	0.988	0.996	0.993	0.989

Standard errors clustered by month in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Notes: Regressions with 35 countries over different horizons. All regressions include a fund-country fixed effect and a country-month fixed effect. The constant is included in the regressions but not shown.

is mainly the magnitude of the coefficient on the expected excess return that is affected.

Table D3 tests the benchmark equations with different β's. Column (1) is the benchmark with $\beta = 0.97$ and the corresponding $\delta = 0.89$. Column (2) and (3) use values of $\beta = 0.95$ and $\beta = 0.99$, which correspond to $\delta = 0.87$ and $\delta = 0.91$.

Finally, Table D4 tests the benchmark equation using fund-specific time-varying weights $z_{i,m,t,-n}$ instead of average weights $\bar{z}_{i,m,t,-n}$ to construct $E_t er^{\delta}_{i,n,t,t+k}$. Equation (12) of the paper becomes

$$E_t \tilde{er}^{\delta}_{i,n,t,t+k} = E_t \sum_{s=1}^{k} \delta^{s-1} er_{n,t+s} - \sum_{m \neq n} z_{i,m,t,-n} E_t \sum_{s=1}^{k} \delta^{s-1} er_{m,t+s} \quad \text{(D.1)}$$

Column (1)-(3) correspond to columns (1)-(3) of Table 3 from the paper when we use $E_t \tilde{er}^{\delta}_{i,n,t,t+k}$ instead of $E_t er^{\delta}_{i,n,t,t+k}$. Results are similar to those in Table 3, but with smaller expected excess return coefficients in columns (2) and (3).

Table D2: Portfolio Regressions, Different Predictions

	No DY	No EP	3 Lags	Interest Rates	Non-Recursive
	(1)	(2)	(3)	(4)	(5)
$(z_{i,n,t-1} + z^{bh}_{i,n,t})/2$	0.916***	0.916***	0.916***	0.910***	0.918***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
$(z_{i,n,t-1} - z^{bh}_{i,n,t})$	0.314***	0.313***	0.313***	0.330***	0.313***
	(0.069)	(0.069)	(0.069)	(0.062)	(0.070)
$E_t er^{0.89}_{i,n,t,t+24}$	3.114***	2.597***	2.034***	1.670***	1.910***
	(0.412)	(0.314)	(0.263)	(0.271)	(0.273)
Observations	316732	316732	316426	300070	316732
R^2	0.988	0.988	0.988	0.988	0.988

Standard errors clustered by month in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Notes: Regressions with 35 countries over the horizon 2002:01-2016:07. All regressions include a fund-country fixed effect and a country-month fixed effect. The constant is included in the regressions but not shown.

Table D3: Portfolio Regressions, Different β's and δ's

	$\beta = 0.97, \delta = 0.9$	$\beta = 0.95, \delta = 0.87$	$\beta = 0.99, \delta = 0.91$
	(1)	(2)	(3)
$(z_{i,n,t-1} + z^{bh}_{i,n,t})/2$	0.916***	0.916***	0.916***
	(0.005)	(0.005)	(0.005)
$(z_{i,n,t-1} - z^{bh}_{i,n,t})$	0.313***	0.313***	0.313***
	(0.069)	(0.069)	(0.069)
$E_t er^{0.89}_{i,n,t,t+24}$	2.324***		
	(0.291)		
$E_t er^{0.87}_{i,n,t,t+24}$		2.770***	
		(0.337)	
$E_t er^{0.91}_{i,n,t,t+24}$			1.906***
			(0.246)
Observations	316732	316732	316732
R^2	0.988	0.988	0.988

Standard errors clustered by month in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Notes: Regressions with 35 countries over the horizon 2002:01-2016:07. All regressions include a fund-country fixed effect and a country-month fixed effect. The constant is included in the regressions but not shown.

Table D4: Portfolio Regressions, Benchmark with Contemporaneous Weights $z_{i,m,-n,t}$

	(1)	(2)	(3)
$(z_{i,n,t-1} + z^{bh}_{i,n,t})/2$	0.928***	0.917***	0.919***
	(0.004)	(0.005)	(0.007)
$(z_{i,n,t-1} - z^{bh}_{i,n,t})$	0.172*	0.313***	0.339***
	(0.090)	(0.069)	(0.068)
$E_t \tilde{er}^{0.89}_{i,n,t,t+24}$	1.033***	1.979***	
	(0.144)	(0.294)	
$E_t \tilde{er}^{0.89}_{i,n,t,t+60}$			3.861***
			(0.849)
Fund-Country FE	Yes	Yes	Yes
Country-Month FE	No	Yes	Yes
Observations	316732	316732	196828
R^2	0.987	0.988	0.990

Standard errors clustered by month in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Notes: Regressions with 35 countries over the horizon 2002:01-2016:07. The constant is included in the regressions but not shown.