Accepted Manuscript

Predicting international stock returns with conditional price-to-fundamental ratios

Jochen Lawrenz, Josef Zorn

PII: S0927-5398(17)30057-9
DOI: http://dx.doi.org/10.1016/j.jempfin.2017.06.003
Reference: EMPFIN 983

To appear in: *Journal of Empirical Finance*

Received date : 13 September 2016
Revised date : 30 May 2017
Accepted date : 25 June 2017

Predicting International Stock Returns with Conditional Price-to-Fundamental Ratios

Jochen Lawrenz[a] Josef Zorn[b]

June 26, 2017

Abstract

Taking the perspective of international asset allocation, this paper tests if predictive regressions conditional on time-series and cross-sectional information can improve forecasts of stock index returns. We use different current price-to-fundamental ratios as predictors and condition the sample on the indicator if time-series and cross-section deliver consistent versus opposing signals. Using panel regressions, we find that only consistent ratios (i) display significant mean-reverting behavior, (ii) provide strong in-sample as well as out-of-sample evidence for return predictability, and (iii) yield economic gains in a Bayesian asset allocation framework.

JEL classification: G11, G12, G15, G17.

Keywords: International stock market returns, international asset allocation, predictability, price to fundamental ratios.

[a]University of Innsbruck, Department of Banking and Finance, jochen.lawrenz@uibk.ac.at.
[b]University of Innsbruck, Department of Banking and Finance, josef.zorn@uibk.ac.at.

1. Introduction

Practitioners in international asset allocation commonly pay attention to the valuation levels of local markets relative to their long-run average as well as relative to a global average. The idea is that countries whose current valuation levels are below (above) the global mean appear to be cheap (expensive) in the cross-section, while countries with valuation levels below (above) their own long-run averages appear cheap (expensive) in the time-series dimension. Economically, this perspective obviously touches on international stock return predictability. More specifically, it raises the research question, if return predictability can be improved by combining the information from the time-series and the cross-section. A long-standing debate discusses predictability of stock returns from valuation ratios and numerous further predictors within domestic markets. A growing literature extends the analysis also to an international perspective. This paper adds to the international predictability evidence by including a country's price-to-fundamental ratio[1] *relative* to its domestic long-run mean as well as *relative* to the global average. Thus, we ask: Does forecastability strengthen when predictive regressions are augmented by including the combined information from the time-series and the cross-section dimension? Intuitively, we might expect that a domestic market which appears expensive not only in the time-series but also in the cross-section dimension provides stronger evidence for predictability. Thus, we condition our sample and the predictive regressions on the indicator if the current local price-to-fundamental ratio is above (below) its own long-run average *as well as* above (below) the long-run global mean. If the signal from both the time-series and the cross-section points in the same direction, we label it a *consistent pe* ratio, while for conflicting signals we call it an *opposing pe* ratio. Our main contribution is a detailed analysis if predictive regressions conditional on consistent versus opposing signals improve return forecasting in- and out-of-sample. We illustrate our approach with the *pe* ratio as a leading example. Importantly though, results are robust against using other predictor variables such as the dividend yield or a price to cash flow ratio.

Essentially, our analysis is a test whether the information from the cross-section holds some additional predictive power. In this sense, our approach is related to contributions by Harvey (1991), Campbell and Hamao (1992), and more recently Rapach, Strauss, and Zhou (2013). In particular, the latter show that information from US markets is a significant predictor for non-US markets. Contrary to other studies from the international predictability literature which basically perform predictive regressions on a dataset extended to global markets,

[1] As price-to-fundamental ratios we use three classical variables: (i) Price-earnings (*pe*), (ii) price-dividends (*pd*) and (iii) price-to-cash flow (*pc*) ratios. As leading variable, we use *pe* ratios, and show that all results also hold for the two alternative valuation ratios.

Harvey (1991), Campbell and Hamao (1992), and Rapach et al. (2013) turn their attention to the issue of intertwined information linkages. Our study is related to this focus except for taking the US market to be pivotal. Instead, we analyze in a more general sense, if information from outside a local market can improve the predictability within this market. From a general perspective, our approach departs from usual unconstrained, linear predictive regressions by conditioning on an economically motivated constraint that makes the predictive regression effectively non-linear. Imposing additional constraints has been shown to improve predictive accuracy. Early contributions are credited to Campbell and Thompson (2008) and Rapach, Strauss, and Zhou (2010) who show that once theoretically and economically motivated sign restrictions are introduced, forecastability and out-of-sample tests are clearly improved. More recently, Pettenuzzo, Timmermann, and Valkanov (2014) expand the approach of Campbell and Thompson (2008) to allow for more efficient use of constraints in the inference of forecasts. Our contribution has a share to this strand of the literature not in the sense of a sign or size restriction of coefficient estimates, but in the sense of constraining the sample to be used in the predictive regression. Importantly, both approaches effectively work similar when it comes to out-of-sample validation.

The analysis is structured in four major parts. First, the belief that the orientation towards a benchmark contains information implicitly assumes some degree of stationarity. Thus, as preliminary evidence towards predictability, we estimate a partial adjustment model. This test is motivated by the possibly mean-reverting behavior of price-to-fundamental ratios which is suggested by the mixed evidence from stationarity tests. We condition the sample on our indicator for consistent versus opposing price-to-fundamental ratios, and find strong support for differences in the partial adjustment behavior. Significant adjustment behavior is only detected for countries with consistent signals from the cross-section and the time-series. Robustness results from a simulation exercise show that these results are neither mechanical nor tautological. This evidence encourages us to proceed to the second step, which consists of in-sample tests for predictive accuracy of the price-earnings ratio. As a major result, we find strong and robust in-sample support for differences in the predictive accuracy. Conditional on the signal from the cross-section and time-series, we find strong predictability for countries with consistent signals, while no significant results are obtained for countries with opposing signals. The evidence is supported from VAR estimations, impulse-response functions as well as from a simulation exercise. The results are in line with the findings from the partial adjustment model in the previous step, and are robust against a variety of econometric concerns. In the third step, we conduct out-of-sample tests. In line with similar studies in the literature, we find modest, albeit slightly positive results for out-of-sample tests in the unconditional setup. However, conditioning the sample on consistent versus opposing price-to-fundamental ratios once again results in notable differences. Importantly, we can confirm the superior in-sample performance of consistent ratios also within out-of-sample tests. While predictions from opposing ratios fail to beat the prevailing average forecast, we find strong positive out-of-sample R^2 for consistent price-to-fundamental ratios, confirming the superior forecasting accuracy also in a real-time setting. In order to assess the practical usefulness, we finally conduct an analysis on how the econometrical evidence of predictability can be transformed to economic gains within an investment context. We incorporate return predictions from out-of-sample models into a Bayesian asset allocation process. We find that the econometrical out-of-sample evidence turns into a significant and sizeable improvement in the economic investment performance. Thus, in a nutshell, we find that the practitioners' intuition seems to contain some economic significance. The information from the time-series as well as the cross-section delivers a significant improvement in in-sample as well as in out-of-sample predictive regressions. Furthermore, we find that the out-of-sample evidence can actually be turned into a significant and sizeable investment gain in a Bayesian asset allocation process.

The remainder of the article is structured in the following way. The next section gives a comprehensive outline on the methodology and the related literature. Section 3 presents the data. Section 4 highlights results for the in-sample analysis of partial adjustment and predictive regressions. Section 5 and 6 add tests for out-of-sample and investment strategies. Section 7 provides additional robustness checks including other predictor variables. Section 8 concludes.

2. Methodology and Related Literature

Stock return predictability has spurred a huge literature with numerous different aspects, one of them being the empirical evidence on the international level. Early contributions in this respect are due to Cutler, Poterba, and Summers (1991), Harvey (1991), Bekaert and Hodrick (1992), Campbell and Hamao (1992), Solnik (1993), Ferson and Harvey (1993), or Richards (1995). More recent evidence is provided by Rapach, Wohar, and Rangvid (2005), Ang and Bekaert (2007), Hjalmarsson (2010), Rapach et al. (2013) or Bollerslev, Marrone, Xu, and Zhou (2014). The focus of one strand of these contributions is to assess if predictability can be considered as a universal feature of developed capital markets. Contrary to that, another strand adopts the perspective of international asset allocation and analyzes to what extent individual local markets are driven by global (macro-) economic factors or to detect cross-country predictive variables. In this respect, early contributions such as Harvey (1991) test excess returns on the world equity market index as predictor, while Ferson and Harvey (1993) use numerous additional macroeconomic variables like global industrial production and global inflation. Campbell and Hamao (1992) focus on the links between the US and Japanese economy and test among others if the US dividend-price ratio can predict Japanese stock returns and vice versa with the intention to detect changing intensities of cross-country market integration. More recently, Rapach et al. (2013) emphasize the pivotal role of the US market by showing that lagged US market returns are a strong predictor of stock returns in non-US industrialized countries. Our contribution fits within the latter strand in the sense that our interest is on the predictive performance conditional on information from global financial markets. However, contrary to the dominant approach in the previous contributions, which is to use fundamental information about the global economy, we adopt a perspective which is motivated from practical decision-making in international asset allocation. A predominant criterion for asset allocation is the orientation towards a relevant *benchmark*. For aggregate international asset allocation, two such relevant benchmarks are the orientation towards the country's own historical average and the orientation towards the global average, i.e. time-series and cross-sectional information. A prominent example with significant outreach among investment practitioners is J.P. Morgan's 'Guide to the Markets'.[2] Within the section on global equity valuations, this investment guide provides information from an aggregate index composed of *pe* ratio, market-to-book ratio, and dividend yield if the current valuation levels of national markets appear cheap or expensive relative to their own 10-year average valuation level and relative to the world average valuation level.[3] Using the information from a multi-year average is basically a technical strategy, which is usually not in the focus of the academic predictability literature as it appears at odds with the idea of efficient markets. However, contributions from Han, Yang, and Zhou (2013) and Neely, Rapach, Tu, and Zhou (2014) demonstrate that technical indicators derived from moving averages, momentum or trading volume do significantly improve the prediction of equity risk premia.

Clearly, the notion that current valuation levels above averages indicate overvaluation implies that price-multiples display some degree of stationarity. Only expecting valuation levels not to persist justifies the belief that a market is over-/undervalued. Figure 1 illustrates our basic research idea graphically.

[2]J.P. Morgan's 'Guide to the Markets is published quarterly and covers a wide range of fundamental and technical economic indicators for global investment decisions. See J.P. Morgan Institutional Asset Management at `am.jpmorgan.com/us/institutional/home`.

[3]See e.g. J.P. Morgan's Guide to the Market, US, 2Q 2015, p. 52, available at `am.jpmorgan.com/gi/getdoc/1383191101905`.

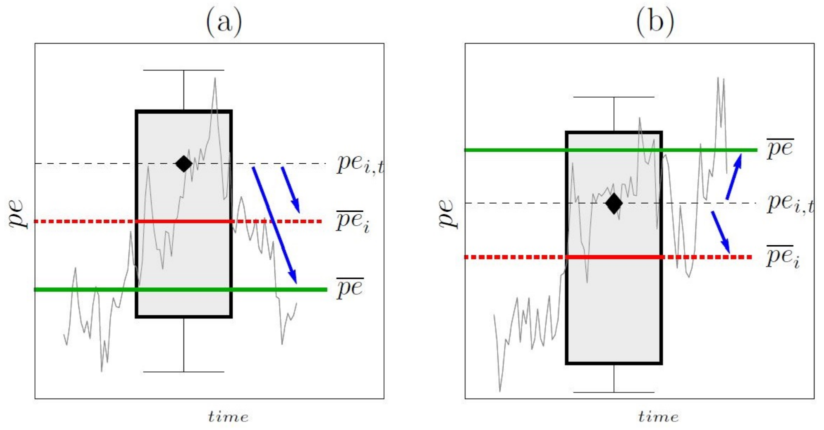

Fig. 1. Schematic display of consistent versus opposing signals

This figure illustrates the research idea. Panel (a) schematically shows the case where the current ratio ($pe_{i,t}$) is above both the domestic mean $\overline{pe}_i$ and the global mean $\overline{pe}$ (consistent signal). Panel (b) shows a case where the current ratio $pe_{i,t}$ is between both means (opposing signal).

Both panels schematically display a boxplot of historical pe ratios for a local market. The (black) diamond locates the current valuation level (light dashed line, indicated by $pe_{i,t}$), with a fictitious sample path of the pe ratio displayed in the background. The thick, dotted (red) line labeled $\overline{pe}_i$ represents the historical within-country average, while the thick, solid (green) line labeled $\overline{pe}$ indicates the global average. In panel (a), the current valuation level $pe_{i,t}$ is above the country mean $\overline{pe}_i$, and the country mean is above the global average $\overline{pe}$, i.e. $pe_{i,t} > \overline{pe}_i > \overline{pe}$. Contrary to that, panel (b) displays a situation where again the current country valuation level $pe_{i,t}$ is above its own historical average $\overline{pe}_i$, but the country average is *below* the global average, i.e. we have $\overline{pe} > pe_{i,t} > \overline{pe}_i$. We group pe ratios according to this distinction into two categories, and call current pe ratios which fit to the description in panel (b),[4] *opposing signal pe* ratios, while pe ratios corresponding to the description from panel (a),[5] are labeled *consistent signal pe* ratios.

The point of distinguishing between the two categories follows from the observation that *consistent signal pe* ratios convey an unambiguous interpretation as the orientation towards both the domestic as well as the global benchmark indicate that the current valuation level is either high ($pe_{i,t} > \overline{pe}_i > \overline{pe}$) or low ($\overline{pe} > \overline{pe}_i > pe_{i,t}$). In contrast, current valuation levels which we group as *opposing signal pe* ratio convey mixed information: On the one hand, the market seems overvalued when assessed against its own history ($pe_{i,t} > \overline{pe}_i$), but on the

[4]Including the reverse situation $\overline{pe}_i > pe_{i,t} > \overline{pe}$

[5]I.e. ratios for which $pe_{i,t} > \overline{pe}_i > \overline{pe}$ and $\overline{pe} > \overline{pe}_i > pe_{i,t}$ hold true

other hand the market seems undervalued against the global benchmark ($\overline{pe} > pe_{i,t} > \overline{pe}_i$). If we believe that the orientation towards the benchmarks holds some information, we should expect stronger effects for *consistent signal pe* ratios than for *opposing signal pe* ratios. As already mentioned, the belief that information from the benchmarks (country and global means) are valuable, implicitly presupposes that *pe* ratios display some stationarity. Actually, there are reasons to expect that price multiples are likely to display some tendencies to revert to reasonable benchmarks, although econometric evidence on stationarity is mixed. If this were true, then mean-reverting tendencies might be present on both, the domestic as well as the global level. These conjectured mean-reverting forces are illustrated in Figure 1 by the corresponding blue arrows pointing towards both mean values. Obviously, *consistent signal pe* ratios distinguish themselves in that the mean-reverting tendencies point in the same direction as opposed to the category of *opposing signal pe* ratios where they point towards opposite directions. Thus, we might expect stronger mean-reverting tendencies for *consistent signal pe* ratios. Although the evidence for mean-reversion is neither necessary nor sufficient to also detect return predictability, it could be taken as an encouraging piece of evidence to pursue predictive regressions. Thus our analysis is structured in such a way as to first find evidence for differences in mean-reverting behavior.

Tests for mean-reverting behavior have a long-standing tradition dating back to e.g. Poterba and Summers (1988), Fama and French (1988), or Bekaert and Hodrick (1992) which detect mean-reverting behavior for equity prices. Fama and French (2000) focus on predictability of earnings and profitability and find mean-reversion from the results of a partial adjustment model. Evidence on a stationary component in valuation ratios is mixed. As discussed in e.g. Lettau and Van Nieuwerburgh (2008), price ratios display surprisingly high persistence. However, they argue that this persistence may be due to structural breaks and once adjusted for regime-shifts, the null of a unit root can be rejected. Given our research objective, we are interested in detecting if *pe* ratios with *consistent signals* display different stationarity characteristics as *pe* ratios with *opposing signals*. However, this intention can hardly be tested by applying the usual stationarity tests conditional on our *consistent-opposing* indicator as it is not a stable property of a time-series. Instead we resort to a method for estimating the speed of adjustment in a partial adjustment model.[6] Using a partial adjustment model seems also justified given that unit root tests are ambiguous for international financial ratios

[6]Partial adjustment models are commonly used in Corporate Finance for the description of dividend policy or leverage dynamics. See e.g. Chen, Da, and Priestley (2012) for dividend policy or Flannery and Rangan (2006) on capital structure. In the context of the predictability literature, partial adjustment models have been applied by e.g. Fama and French (2000) or Ahn, Boudoukh, Richardson, and Whitelaw (2002).

mainly due to low power of the tests.[7] An early application of partial adjustment models to *pe* ratios is due to Lev (1969) who tests for individual firm-level adjustment to a group (i.e. industry) mean. Building upon this, we take the analysis from the firm level to the country level and test whether individual countries' *pe* ratios partially adjust to a country-group (or global) mean. Results from this test provides evidence to what extent we can detect a common long-run mean in valuation ratios, and in particular if adjustment behavior differs conditional on our distinction of *consistent* versus *opposing* signals.

The major part of our contribution revolves around predictive regressions conditional on the *consistent-opposing signal pe* indicator. Stock return predictability is heavily researched since the early contributions of Rozeff (1984), Fama and French (1988) and Campbell and Shiller (1988a,b). While it was first taken as evidence against market efficiency,[8] it is by now seen as reflecting time-varying expected returns which is widely accepted as a "*new fact in finance*" (Cochrane, 1999). Whereas the explanations for time-varying expected returns are manifold, the evidence for long horizon predictability is hard to neglect.[9] Comprehensive reviews on methodology and empirical evidence are offered by Cochrane (2005), Lettau and Ludvigson (2009) or Rapach and Zhou (2013).
A plethora of economic variables have been tried for improving forecast accuracy, giving rise to model uncertainty and data snooping concerns. Cremers (2002) and more recently Schrimpf (2010) and Turner (2015) address this issue by adopting a Bayesian perspective to model selection. They find mixed results depending on the prior weights given to predictors, but conclude that evidence for predictability weakens with model uncertainty. With this disclaimer in mind, we still opt for a predictive regression which focuses on *pe* ratios as a leading example for the following reasons: First, it represents the direct test of the practical investment advice outlined above. Practitioners have for long considered price-earnings ratios as a measure of fundamental value, dating back to e.g. Graham and Dodd (1934). Second, the dividend yield as another prominent predictor variable is criticized due to the dividend disappearance phenomenon and the shift to share repurchases.[10] Third, dividend smoothing as a choice of corporate policy affects predictability (Chen et al., 2012). For these

[7] See for example Ang and Bekaert (2007), Park (2010) or Mcmillan and Wohar (2013)

[8] As e.g. Campbell and Shiller (1988a)

[9] Among the most prominent explanations in the literature are time-varying risk aversion, time-varying aggregate consumption risk, time-varying consumption disasters, time-variation in risk-sharing opportunities among heterogeneous agents or time-variation in beliefs. See e.g. Koijen and Van Nieuwerburgh (2011) for a survey.

[10] See e.g. Fama and French (2001); Campbell and Shiller (2001); Boudoukh, Michaely, Richardson, and Roberts (2007); Kellard, Nankervis, and Papadimitriou (2010)

reasons, we follow the multitude of studies[11] which use the price-earnings ratio as predictor. However, to address data snooping concerns we also use the dividend yield and a price to cash flow ratio as robustness checks.

Besides model uncertainty, model *instability* is another challenge for inference from predictive regressions. Lettau and Van Nieuwerburgh (2008) 'reconcile' mixed predictability evidence by pointing towards potential structural breaks. For an international data set, Paye and Timmermann (2006) show that for the vast majority of countries, instability issues arise and structural breaks are heterogeneous across countries and time. Similarly, Giot and Petitjean (2006) find no common pattern of stock return predictability across countries neither on statistical nor on economic grounds. More recently, Henkel, Martin, and Nardari (2011) show that predictability is diminished greatly during business cycle expansions in developed countries. Similarly, Dangl and Halling (2012) confirm a strong link between predictability and the business cycle, while Mcmillan and Wohar (2013) analyze predictability in a panel setting and reveal that return predictability and dividend growth predictability switches from decade to decade. Our results are consistent to this strand of the literature by showing that the strength of (in-sample) predictability depends on the joint information from the time-series and the cross-section which thereby induces time variation.

The third step of our analysis addresses out-of-sample evidence. It is mostly credited to Goyal and Welch (2008) for being the first comprehensive study which shows that out-of-sample evidence (OOS) for the bulk of predictors is weak at best.[12] While subsequent support for weak OOS evidence was raised in numerous contributions, a backlash came from contributions such as Campbell and Thompson (2008) and Rapach et al. (2010) who show that once theoretically and economically motivated sign restrictions are introduced, forecastability and out-of-sample tests are clearly improved. Furthermore, ongoing debate exists as to whether in-sample or out-of-sample tests are more useful for inference–a point made by e.g. Inoue and Kilian (2005). Apparently, the return predictability evidence is not so clear-cut as it was conceived and the jury is still out.

For our purposes however, tests for OOS evidence are clearly insightful as they arguably resemble most closely the situation of a real-time investor who needs to make decisions without the information set of an econometrician. Thus, we complement and contrast our in-sample evidence with various out-of-sample diagnostics.

[11] See e.g. Hodrick (1992), Lamont (1998) or Lewellen (2004).

[12] They conclude that "*by and large, these models have predicted poorly both in-sample (IS) and out-of-sample (OOS) for 30 years now; these models seem unstable, as diagnosed by their out-of-sample predictions and other statistics; and these models would not have helped an investor with access only to available information to profitably time the market*" (Goyal and Welch, 2008, p.1455).

In contrast to the vast literature on econometric evidence of in- and out-of-sample predictability, only few contributions also pay attention to the question of how (econometrical) predictability turns into (economical) investment performance. As a final section in our contribution, we analyze to what extent the forecastability of returns can be transformed into a profitable investment strategy. A first attempt follows the contribution of Campbell and Thompson (2008) and calculates the utility gain of a mean-variance investor. A more sophisticated approach is following the Bayesian strategy of Black and Litterman (1991, 1992). Using several portfolio performance measures, we test whether the model including both the time-series and cross-sectional information outperforms the standard predictive model.

3. Data

The data is gathered from Thomson Reuters Datastream. Time series are extracted for equity indices of 27 countries and 14 regional equity indices on a monthly frequency and in US$ terms.[13] Datastream provides the total return index that takes into account dividend payments.[14] pe ratios are calculated by dividing market value by the realized total earnings at time t (trailing pe). Since emerging markets exhibit fundamental differences to developed markets with respect to financial market development, return and predictability patterns (see e.g. Harvey 1995), this paper divides the sample into (i) developed markets, (ii) emerging/developing markets and (iii) regions.[15] Table 1 provides summary statistics for returns and pe ratios. Due to data availability the sample ranges from 1973M01 to 2014M05 for developed countries and for regions. For emerging/developing countries the sample ranges from 1991M01 to 2014M05. Series with shorter horizons are excluded and therefore the sample is dictated by data availability. pe ratios exhibit similar moments for most of the countries. Still, Japan can be seen as an outlier in developed markets with a mean

[13] We focus on US$ returns since we are primarily interested in the asset allocation implications of predictability for US investors. Critically, this might include effects of exchange rate fluctuations on returns. Jordan, Vivian, and Wohar (2015) note that if exchange rates follow a random walk the numéraire might only affect the standard error of the coefficient in a predictive regression (in the setting $r = f(pe)$). They show that indeed the choice of currency matters for predictability. However, our results with local currencies (not reported) reveal that this effect is limited and does not change our findings qualitatively. In fact, predictability in local currencies is even stronger.

[14] Datastream computes the total return index as follows,

$$RI_t = RI_{t-1} \frac{PI_t}{PI_{t-1}} \left(1 + \frac{D/P}{100n}\right)$$

where RI_{t-1} is the return index on the previous day, PI_t is the price index on day t, D/P denotes the dividend-price ratio and n is the number of days in financial year.

[15] The division into these three distinct groups is (politically) debatable but for the purpose of this study we find it suitable.

ratio of 35.39. In emerging countries, Taiwan with a mean *pe* ratio of 20.17 and in regions, the Far East with 27.81 display the highest mean values. Group means are very similar for developed and emerging countries, 16.31 and 16.28 respectively. Hongkong, Thailand, Scandinavia and SE (South East) Asia show a distinct high kurtosis. Mean returns are all positive ranging from 0.18 to 1.10 %.

[Insert Table 1 near here]

Stationarity of *pe* ratios is important for estimation of the model since it implies a cointegration relationship between prices and earnings. Table 2 shows individual and panel unit-root tests. The tests already highlight the disparity between the different methodologies and the potential low power of unit-root tests. The first column shows the autocorrelation AC(1) coefficient. Clearly, (log) *pe* ratios are highly persistent with coefficients near 1. The second column provides Augmented Dickey Fuller (ADF) test results (Said and Dickey, 1984). The null of a unit-root can be rejected for 26 out of 41 indices. Conversely, KPSS test statistics (Kwiatkowski, Phillips, Schmidt, and Shin, 1992) highlight that the null of stationarity can be rejected for all but 7 countries. The Phillips-Perron test (Phillips and Perron, 1988) of the null of a unit-root can be rejected for 25 out of 41 countries/regions. Since we are interested in the panel properties the panel unit root test of Levin, Lin, and Chu (2002) is conducted. The null hypothesis of a unit-root can not be rejected for developed countries and regions. However, the Im, Pesaran, and Shin (2003) test which assumes individual unit root processes provides strong evidence for the rejection of a unit-root process. These ambiguous results are common in the literature and can be primarily associated with the low power of the test statistics, small sample sizes, long memory processes, non-linearity in the cointegration relationship or structural breaks.[16] For these reasons, we provide results from a partial adjustment model to find evidence whether *pe* ratios do revert to a specific mean.

[Insert Table 2 near here]

4. In-Sample Analysis

4.1. Partial adjustment

We first determine whether *pe* ratios adjust to a global mean ($\overline{pe}$) between countries by employing a partial adjustment model. Lev (1969) provides evidence for partial adjustment

[16]See for example Lamont (1998), Koustas and Serletis (2005) , Campbell and Yogo (2006), Nagayasu (2007), Lettau and Van Nieuwerburgh (2008) or Park (2010).

of pe ratios across companies. We extend this analysis to cross-country ratios in an asset pricing context related to Fama and French (2000) or Chen et al. (2012).

Consider the following partial adjustment model,

$$pe_{i,t+k} - pe_{i,t} = \lambda_k \left(\overline{pe} - pe_{i,t}\right) + \varepsilon_{i,t+k}, \quad 0 < \lambda_k \leq 1, \tag{1}$$

or equivalently,

$$pe_{i,t+k} = \overline{pe}\lambda_k + (1 - \lambda_k)\, pe_{i,t} + \varepsilon_{i,t+k}, \tag{2}$$

where $pe_{i,t+k}$ denotes the pe ratio for country i at time $t + k$. k is the horizon for which the partial adjustment is considered (1 and 5 years). $\overline{pe}$ is the mean pe ratio across countries or regions. λ_k denotes the speed of adjustment parameter which lies between 0 and 1.

Table 3 presents results for the partial adjustment model (Eq. 2) conditional on the *consistent-opposing signal* indicator. λ measures the speed of reversion. If $\lambda = 1$, the pe ratio fully adjusts to the target ratio $\overline{pe}$ within the period. A common way to gain intuition for the magnitude of λ is to translate them into half-lives. A half-live denotes the time it takes a country's pe ratio to adjust back one half the distance to the mean after a one unit shock to the error term $\varepsilon_{i,t+k}$. For an AR(1) process half-lives are computed $log(0.5)/log(1 - \lambda)$. A λ of 0.256 for the unconditional case (see column 2) implies a half live of 2.34 years for the sample of developed countries. Japan is commonly seen as exceptional case due to high pe ratios. For the sample of developed countries excluding Japan the speed of adjustment increases to 0.319 and an implied half-live of 1.80 years. For emerging countries and regions the implied half live is 0.93 years and 3.69 years respectively.

[Insert Table 3 near here]

Splitting the sample according to our opposing-consistent indicator offers interesting findings: Column 3 and 4 in Table 3, which are labeled by $\hat{\lambda}_o$ and $\hat{\lambda}_c$, report adjustment speeds conditional on opposing and consistent signals respectively. It turns out that the two samples display different patterns. In particular, adjustment speeds for opposing signals ($\widehat{\lambda}_o$) are not significant for the three groups except 1 and 5 years for emerging countries. Even worse, adjustment speeds show wrong (negative) signs. For all other samples adjustment speeds are highly significant. Slightly higher coefficients across the three groups are shown for consistent signals ($\widehat{\lambda}_c$), suggesting higher adjustment speeds to the group mean when pe ratios are far distant. Critically, these results do not change qualitatively when using the median as a metric for global and country averages. For each country or regional time-series there are several periods where pe ratios display opposing signals. See Section B in the Appendix for summary statistics and a visual description of opposing versus consistent signals.

For a couple of reasons, one may conjecture that the mean-reverting behavior has changed over the observation period. Figure 2 depicts the speed of adjustment for the three groups through time using a rolling window of 5 years. In the aftermath of the oil crisis, adjustments were practically zero. Particularly in the most recent crisis, adjustment speeds spiked from around 0.2 to 0.8 for developed and emerging countries. Overall, emerging countries adjust faster to their mean *pe* ratio. Visually, from the 1990s on, adjustment speeds seem to exhibit a positive trend, which may be due to increased market integration. A finding which is consistent with results from Campbell and Hamao (1992).
Besides looking more closely at the temporal development, one could also investigate in more detail cross-sectional differences in the mean-reversion behavior. Table 3 distinguishes between three samples and shows different adjustment speeds. Due to our panel approach, we cannot provide country-specific estimates, which could have delivered interesting insights into differential cross-country behavior. Although being interesting in itself, we do not pursue a more detailed analysis of the temporal behavior or cross-sectional differences of the mean-reversion since it is not our primary research focus in this contribution. Hence, we turn to in-sample predictive regressions in the next section.

[Insert Figure 2 near here]

4.2. Return predictability

Campbell and Shiller (1988a,b) link financial ratios to expected stock returns. Their dynamic dividend discount model is the main underlying framework for the return predictability literature. The model implies that time-varying financial ratios must predict either expected returns or dividend growth or both by definition of the identity (see e.g. Cochrane, 2008).

The well known approximate identity (Campbell and Shiller, 1988a,b) reads

$$p_t - d_t = const. + \sum_{j=1}^{\infty} \rho^{j-1}(\Delta d_{t+j} - r_{t+j}), \tag{3}$$

where lower case letters emphasize logarithms and $p_t - d_t$ is the price-dividend ratio. Δd_{t+j} denotes dividend growth and r_{t+j} future returns.[17]

[17] The model relies on the property that $p_t - d_t$ does not explode faster than ρ^{-t}, $\lim_{j\to\infty} \rho^j(p_{t+j}+d_{t+j}) = 0$. Thus, rational bubbles are ruled out in this approximation. The non-explosiveness is somewhat controversial in the literature. For instance, Lewellen (2004) and Cochrane (2008) rule out rational bubbles because, economically, $d-p$ should be stationary since dividends and prices are cointegrated with the vector $[1, -1]'$. A unit root would mean that the series is unbounded and could take negative values and go to infinity

Eq. (3) holds ex post. Taking conditional expectations relates pd ratios to ex ante dividend growth and return forecasts. The interpretation of this accounting identity is straightforward. High pd ratios must be followed by high dividend growth Δd_{t+j} or low returns r_{t+j} or a combination of both.

Following Chen et al. (2012) we can rewrite the Campbell-Shiller model for the price earnings ratio,

$$p_t - e_t = const. + \mathbb{E}_t\left[\sum_{j=1}^{\infty} \rho^{j-1}(\Delta e_{t+j} + (1-\rho) de_{t+j})\right] - \mathbb{E}_t\left[\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}\right], \tag{4}$$

where de denotes the log payout ratio $DE_t = D_t/E_t$. To check for potential differences in our analysis with respect to payout policy we use the dividend yield and the price to cash flow ratio as additional predictor variables.

The Campbell-Shiller model gives rise to empirical testing via time-series predictive regressions. We advance this framework to panel econometrics following Hjalmarsson (2010) since we focus on cross country predictability. Consider the following pooled panel data model,

$$r_{i,t\to t+k} = \alpha_k + pe_{i,t}'\,\beta_k + u_{i,t\to t+k}, \qquad \forall i = 1,\ldots,n, \quad t = 1,\ldots,T, \tag{5}$$

$$u_{i,t\to t+k} = \phi_k\, u_{i,t} + \varepsilon_{i,t\to t+k}, \tag{6}$$

where $r_{i,t\to t+k}$ denotes the cumulated total return from period t to $t+k$ for country i. α_k is the common intercept and β_k the predictive coefficient for (log) pe ratios $pe_{i,t}$. Although most studies in the return predictability literature find it sufficient to add robust standard errors in order to address issues with autocorrelation, we follow a more conservative approach. Particularly, since $u_{i,t\to t+k}$ exhibits autocorrelation rising with the horizon, we include an AR(1) term for inference with overlapping observations. Thus, by modeling the error term with an AR(1), we achieve $\varepsilon_{i,t\to t+k} \sim (0,\Sigma)$.

Predictive regressions are subject to several critiques. First, overlapping observations in long-horizon regressions result in serially correlated fitted residuals. Valkanov (2003) stresses that this problem arises particularly when the return horizon k is large compared to the sample size T. Second, persistent, predetermined regressors result in upwards biased predictor coefficient estimates (Stambaugh, 1999). However, as Hjalmarsson (2010) points out, in a pooled panel model the estimator is unbiased as long as no fixed effects are included.

which is counterintuitive. However, Engsted and Pedersen (2010) provide a bubble model that explains return predictability. Still, the authors find that the approximation error in the Campbell-Shiller model is surprisingly small even in presence of explosive bubbles. Koustas and Serletis (2005) test for long memory properties in the price-dividend ratio, suggesting a fractional cointegration relationship of order < 1 which is inconsistent with the existence of rational bubbles.

Thus, the Stambaugh (1999) bias is diminished since cross-sectional information mitigates the endogeneity effects.

Noteworthy, Ang and Bekaert (2007) find that statistical inference at long horizon regressions depends critically on the choice of standard errors. To correct for heteroskedasticity and serial correlation in panels we use the Panel Corrected Standard Error (PCSE) methodology (Beck and Katz, 1995) that uses Seemingly Unrelated Regressions (SUR). Ang and Bekaert (2007) show that this methodology is a generalization of Hodrick (1992) to panel regressions. We prefer these standard errors to the Newey and West (1987) methodology since they are more conservative.

An important question is whether the specification of a common intercept α_k and a common slope coefficient β_k is economically and statistically justified. Practically, even when the individual coefficients β_k are not identical, the pooled estimate can be interpreted as the average predictive relationship across countries. Economic theory provides some guidance concerning fixed effects. A common intercept of $\alpha_i = \alpha$ for all i enforces that all countries have the same expected return in the case of no predictability, $\beta_i = 0$. This is not the case as empirical studies for equity premiums across countries demonstrate (see e.g. Jorion and Goetzmann, 1999). Considering the capital asset pricing model (CAPM), a common intercept would imply that CAPM betas are the same for all countries. Again, this commonality can be rejected since CAPM betas vary substantially across countries and are even negative for some emerging markets (Harvey, 1995). Hjalmarsson (2010) concludes that from economic considerations, fixed effects should be included. However, his results indicate that pooling the data is often justified since the null of a common slope coefficient cannot be rejected in statistical tests. In light of this discrepancy and to address this issue we analyze developed and emerging countries separately to counter large economic deviations in coefficients.

Another important empirical test in light of Cochrane (2008, 2011) elaborates whether earnings growth is predictable. The absence of earnings growth predictability is further and arguably even stronger evidence for return predictability. We address this issue within a VAR model and by impulse-response analysis.

The evidence for return predictability is primarily measured by the R^2 statistic. This measure tends to be higher at long horizons which is seen as further evidence for predictability (Cochrane, 2008). However, Boudoukh, Richardson, and Whitelaw (2008) argue that short and long horizon estimates are highly correlated, providing little information advantages. Koijen and Van Nieuwerburgh (2011) elaborate on both views and arrive at the conclusion that R^2 first rises with horizon before it decreases. Cochrane (2008) finds power advantages arising already beyond 5 years. We test horizons from 1 month up to 5 years and correct for correlation issues.

Table 4 summarizes our main results for the pooled predictive regression (Eq. 5) in the unconditional setting as well as conditional on our indicator for consistent versus opposing *pe* ratios. The first two columns show coefficients with standard errors and R^2 for the whole sample (*all*). Across all three panels, coefficients have the correct negative sign. High current *pe* ratios are associated with lower future returns. The coefficients are even highly significant as emphasized by panel heteroscedasticity and serial correlation robust (period SUR) standard errors.[18] To account for overlapping observations, we additionally infer significance from the AR(1) model as highlighted by the asterisks. Both, coefficients as well as R^2 rise with horizon. From an economic perspective, R^2 values are huge for 5 year horizons, reaching 25.2% for emerging countries.

Now, if we add our conditioning information, we obtain interesting insights: For *pe* ratios with opposing signals, i.e. ratios between the group mean and the country/region mean, the picture looks entirely different. Five coefficients change the sign, none are significant, and R^2 is very low, particularly for emerging countries. Apparently, returns are not predictable when signals are opposing, thereby mirroring the results from the partial adjustment model.

In contrast, for the sample with consistent signals, predictability is even stronger compared to the whole sample. All coefficients are highly significant and R^2 reaches 27.4% for emerging countries. Strong predictability is obtained for all samples. By looking at samples above both means we see considerable higher R^2 compared to those above only the country mean. The same holds for sample below both means. The differences emphasize the importance of both, the time-series mean ($\overline{pe_i}$) and the cross-sectional mean ($\overline{pe}$). However, better predictive results are achieved for samples above both means ($> \overline{pe}$, $> \overline{pe_i}$) compared to those below both means.

[Insert Table 4 near here]

Common test statistics support our assumption about the pooled model using a common intercept.[19] Importantly though, using the fixed-effects model as a robustness check, does not change the results qualitatively (results not reported).

Additionally to panel regressions we run country specific predictive regressions to evaluate

[18] We also applied Newey and West (1987) standard errors which are even smaller. However, we stick to the PCSE SUR method since it is more conservative.

[19] Tests for redundant fixed effects yield the following results for the cross-section F-stat (sum-of-squares method): 0.0948, 0.3256 and 0.0677 for developed, emerging and regional indices respectively. The χ^2 statistic (likelihood function) yields 1.32942, 3.5940 and 0.8813. All these results indicate that the null hypothesis, H_0:Effects are redundant, cannot be rejected and therefore support the assumption of pooled panel regressions (common intercept).

the predictability for each country/region separately by the following model,

$$r_{t \to t+k} = \alpha_k + \beta_k pe_t + \varepsilon_{t \to t+k}. \tag{7}$$

The intuition and interpretation is similar to the panel case above. Yet, further careful attention must be taken with respect to the Stambaugh (1999) bias. For this reason the point estimates and standard errors are calculated following the procedure suggested by Amihud and Hurvich (2004).

Table 5 reports individual predictive regressions for all countries. Similar to previous research, return predictability differs across countries. However, most coefficients show the expected negative sign except for Belgium (1 and 3 month horizon), Malaysia (1 and 3 month horizon) and regions EMU, Far East, NORCS, Pacific Basin and PIIGS (each 1 month horizon). Out of 41 indices, 7 are highly significant for the 1 month horizon, 6 are highly significant for the 3 month horizon and 8 for the 6 month horizon. R^2 rise mechanically with horizon for virtually all countries, as emphasized by Boudoukh et al. (2008). For this reason we base inference mainly on the 1 month horizon. Typically, US, Hongkong, UK and Singapore show significant return predictability at the 1 month horizon with small R^2 between 2.3% and 0.6%. For developing countries, Chile, India and South Africa display highly significant return predictability. Across regions, the Americas, North America and SE Asia are significant. Overall, the individual results display a rather heterogeneous picture, mirroring results of previous international studies as for instance in Giot and Petitjean (2011).

[Insert Table 5 near here]

4.3. Simulation in VAR

Cochrane (2008) emphasizes the return predictability evidence implied by a first order VAR. We follow this approach and, neglecting payout policy for a moment, we simply substitute earnings for dividends. The approximate identity $r_{t+1} = \rho(p_{t+1} - e_{t+1}) + \Delta e_{t+1} - (p_t - e_t)$ implies that coefficients in a first order VAR representation,

$$r_{t+1} = a_r + b_r(e_t - p_t) + \varepsilon^r_{t+1}, \tag{8}$$

$$\Delta e_{t+1} = a_e + b_e(e_t - p_t) + \varepsilon^e_{t+1}, \tag{9}$$

$$e_{t+1} - p_{t+1} = a_{pe} + \phi(e_t - p_t) + \varepsilon^{ep}_{t+1}, \tag{10}$$

obey the following parameter restriction

$$b_r = 1 - \rho\phi + b_e, \tag{11}$$

where b_r, $\rho\phi$ and b_e are predictive coefficients for returns, pe ratio and earnings growth respectively. They sum up to one and thus we can infer coefficients as well as error terms of one VAR equation from the other two. This implies that the absence of earnings (dividend) growth predictability gives even stronger evidence for the presence of return predictability (Cochrane, 2008). Complementary, Van Binsbergen and Koijen (2010) and Ferreira and Santa-Clara (2011) show that apart from a low frequency component, earnings growth is nearly unpredictable.

The purpose of this exercise is to evaluate return predictability dynamically with a VAR based on Eqs. (8), (9) and (10). The advantage of this approach is that we account for potential earnings predictability and autocorrelation in the system. Through impulse-response functions the effects of shocks to pe ratios are elaborated across time. The functions highlight how earnings and returns react to shocks of pe ratios. For convenience, we suppress subscripts i since data for each variable are stacked including all countries i for the three regions.

In matrix form, a first order VAR representation can be expressed as

$$\begin{bmatrix} r_{t+1} \\ \Delta e_{t+1} \\ e_{t+1} - p_{t+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & b_r \\ 0 & 0 & b_e \\ 0 & 0 & \phi \end{bmatrix} \begin{bmatrix} r_t \\ \Delta e_t \\ e_t - p_t \end{bmatrix} + \begin{bmatrix} \epsilon^r_{t+1} \\ \epsilon^d_{t+1} \\ \epsilon^{ep}_{i,t+1} \end{bmatrix},$$

or in compact parsimonious notation,

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \boldsymbol{\varepsilon}_{t+1}, \tag{12}$$

where $\mathbf{x}_t$ is the vector of endogenous variables $\{r_t, \Delta e_t, e_t - p_t\}$. $\mathbf{A}$ is the matrix representing the time invariant coefficients for lagged endogenous variables. $\boldsymbol{\varepsilon}_{t+1}$ labels the vector of residuals with the property $\boldsymbol{\varepsilon}_{t+1} \sim (0, \Sigma_\varepsilon)$. Given the high persistence of pe ratios, an important characteristic of a VAR(p) process is stability,

$$\det(I_K - A_1 z - \ldots - A_p z^p) \neq 0 \quad \text{for } |z| \leq 1.$$

The VAR is stable if all roots z lie outside the unit circle.

Following Cochrane (2008, 2011) we simulate impulse-response functions using the Cholesky decomposition. Particularly, we show how a shock to the pe ratio affects returns and earnings growth for a 5 year horizon conditional on the current value of the ratio.

Figures 3, 4 and 5 report accumulated impulse-response functions for the three country/regional groups respectively. The functions are plotted for a 5 year response horizon. Standard errors emphasized by the dotted lines are computed via Monte Carlo simulations and 1000 repetitions.

Panels (1a), (2a) and (3a) show responses to one standard deviation innovation to the *pe* ratio affecting the three variables for the whole sample. The *pe* ratio in response rises but the effect diminishes as expected. This emphasizes the mean reverting tendency of *pe* ratios. For returns, the shock shows the correct negative sign but also diminishes through time. For earnings growth the picture is different. A shock affects earnings growth positively but it reverses across time. Thus, in line with results for the dividend yield in Cochrane (2008), earnings growth is not predictable.[20]

Panels (1b), (2b) and (3b) show impulse-response functions for *pe* ratios with opposing signals. Similar to the point estimates in the previous sections shocks to *opposing signal pe* ratios show no particular pattern. Apparently, when the time-series and cross-sectional information oppose, ratios rather follow a random walk. Panels (1c), (2c)and (3c) report results for *pe* ratios with consistent signals. The dynamic effects are slightly more pronounced than the full sample for all three groups.

[Insert Figure 3 near here]
[Insert Figure 4 near here]
[Insert Figure 5 near here]

In general, the paths for the impulse-response functions are in line with the previous results from predictive regressions. Not only are point estimates different across the two subsamples, also the dynamics between the three variables $(pe, \Delta e, r)$ is heavily dependent on the signal from the time-series and the cross-section. The results of the VAR emphasizes the mitigated return dynamics for *pe* ratios with opposing signals.

Importantly, these results are robust to a variety of alternative specifications including different lag lengths, different horizons and even the Cholesky ordering. In general, all specifications are stable with no root z inside or on the unit circle.

[20] We additionally ran predictive regressions for earnings growth, $\Delta e_{i,t \to t+k} = \alpha_k + pe'_{i,t}\ \beta_k + \varepsilon_{i,t \to t+k}$, finding no significant coefficients β_k irrespective of horizon (but using robust standard errors or an AR(1) term).

4.4. *Results for simulated data*

To see whether the differences between consistent versus opposing signals in the partial adjustment framework and the predictability framework are not a tautology, we perform the same tests with simulated data. This exercise should rule out the possibility that our previous results are just an artifact and should shed light to whether the joint time-series and cross-sectional information really alters the results. Since *pe* ratios are assumed to be highly persistent but stationary processes, we simulate data from a stochastic process similar to Ang and Liu (2007).

Consider the continuous-time pendant to the discrete AR(1) process, the Ornstein-Uhlenbeck process with the stochastic differential equation

$$dX_t = \theta(\mu - X_t)\, dt + \sigma\, dW_t, \tag{13}$$

where $\theta > 0$ is the rate by which shocks dissipate, μ denotes the equilibrium mean, $\sigma > 0$ the volatility parameter and dW_t is the increment of a standard Wiener process. The Ornstein-Uhlenbeck process is mean reverting and converges to a stationary distribution.

Integrating from 0 to t with the help of Itô's lemma yields

$$X_t = X_0\ e^{-\theta t} + \mu(1 - e^{-\theta t}) + \int_0^t \sigma e^{\theta(s-t)} dW_s.$$

It can be shown, given X_0 is a constant or a normally distributed variable, that

$$\mathbb{E}[X_t] = X_0\ e^{-\theta t} + \mu(1 - e^{-\theta t})\,,$$

and

$$\mathrm{Cov}[X_s, X_t] = \frac{\sigma^2}{2\theta}\left(e^{-\theta|s-t|} - e^{-\theta(s+t)}\right).$$

This implies that the first moment converges to $\lim_{t\to\infty} \mathbb{E}[X_t] = \mu$, and the variance to $\lim_{t\to\infty} \mathbb{V}[X_t] = \frac{\sigma^2}{2\theta}$. Using this process, we simulate artificial *pe* ratios. However, in order to simulate *pe* ratios close to the empirical ones, we match the model parameters to the empirical estimates. To achieve similar (higher) moments, we extend the model and incorporate different starting values and heterogeneous persistence parameters and different means.

Figure 6 shows the distribution of coefficients λ_k (Eq. 1) for simulated data following the extended model matched to empirical moments. The figure plots the distribution of 100,000 repetitions for 15 artificial *pe* ratios. The dotted line emphasizes partial adjustment speeds for opposing signals whereas the solid line depicts the distribution for consistent signals. Both distributions show roughly the same moments and therefore can hardly explain the huge differences (even negative values) of empirical partial adjustment speeds. For further details of the simulation approach and numerical results see Appendix A. The results are virtually the same for the data generated with the simple Ornstein-Uhlenbeck process. Running predictive regressions with simulated *pe* ratios show no differences when conditioned on signals. Predictive coefficients exhibit the same sample moments for both signal categories. So do impulse-response functions. The results for simulated data indicate that both mean reversion (or partial adjustment) and return predictability should be the same for *pe* ratios with consistent or opposing signals.[21] High persistence and serial correlation do not explain the differences in the previous models. Neither do different sample sizes in the signal groups. The main results are therefore not mechanical nor are they a tautology.

[Insert Figure 6 near here]

5. Out-of-Sample Analysis

Despite an ongoing discussion about whether in-sample or out-of-sample tests are more useful for inference (see e.g. Inoue and Kilian (2005)), Goyal and Welch (2008) provide a thorough discussion of the usefulness of out-of-sample tests. Since predictive regressions are subject to many statistical problems, the inclusion of these tests is important, particularly with respect to possible data mining and model instability aspects. Furthermore, out-of-sample tests are ways to mimic the behavior of real-world investors.

To see whether conditional *pe* ratios (i.e. conditional on consistent or opposing signal) improve out-of-sample predictions, we compare the unconditional forecast with one that includes our consistent versus opposing indicator. Does the forecast strengthen by including this additional piece of information? The two rival forecasts are the unconditional prediction from *pe* ratios which we label as Model 1, following from

$$r_{i,t\to t+k} = \alpha_k + pe_{i,t}'\,\beta_k + \varepsilon_{i,t\to t+k}, \tag{14}$$

and the prediction with the information about the consistent versus opposing indicator, which we label as Model 2, following from

$$r_{i,t\to t+k} = \alpha_k + pe_{i,t}'\,\beta_k + (D_{i,t}^{oos} \times pe_{i,t})'\,\gamma_k + \varepsilon_{i,t\to t+k}. \tag{15}$$

[21] We did simulate additional processes including one following a sine wave which exhibits the highest persistence of a mean reverting process. Again, we found no differences in mean reversion across signal groups.

Importantly, we use only information from previous periods for each forecast. Since our hypothesis following the results of the previous section is to expect stronger predictability for forecasts including the joint time-series and cross-sectional information, we incorporate a slope dummy variable. Sample averages and the dummy variable $D_{i,t}^{oos}$ are calculated using information from previous periods exclusively. Formally, $D_{i,t}^{oos} = 1$ if the pe ratio of country i in time t can be classified as a consistent signal. Each forecast uses an expanding window approach from t_0 to time t.

These arrangements assure that the forecasts are truly out-of-sample. For a forecast comparison, we use several metrics, the Mean Absolute Error, MAE$=\frac{1}{n}\sum_{t=1}^{n}|e_t|$, the Root Mean Squared Error, RMSE$=\sqrt{\frac{1}{n}\sum_{t=1}^{n}e_t^2}$ and the (scale invariant) Theil Inequality Coefficient,

$$TIC = \frac{\sqrt{\frac{1}{n}\sum_{t=1}^{n}(\hat{r}_t - r_t)^2}}{\sqrt{\frac{1}{n}\sum_{t=1}^{n}\hat{r}_t^2} + \sqrt{\frac{1}{n}\sum_{t=1}^{n}r_t^2}}, \qquad 0 < TIC < 1.$$

Complementary, we estimate out-of-sample R^2 following Campbell and Thompson (2008) to see whether both forecasts outperform the historical average,

$$R_{oos}^2 = 1 - \frac{MSFE(\hat{r}_t)}{MSFE(\bar{r}_t)} = 1 - \frac{\sum_{t=T_1+1}^{T}(r_t - \hat{r}_t)^2}{\sum_{t=T_1+1}^{T}(r_t - \bar{r}_t)^2},$$

where $\bar{r}_t$ emphasizes the historical average return for each country in time t.

Diebold and Mariano (1995) provide a statistic to compare two rival forecasts. More precisely, they test the null hypothesis that the two forecasts have the same accuracy by using the following statistic,

$$DM = \frac{\bar{d}}{\sqrt{\frac{2\pi\hat{f}_d(0)}{T}}},$$

where $\bar{d}$ is the sample mean of the loss differential (d_t) between two forecasts. $\hat{f}_d(0)$ is a consistent estimate of $f_d(0)$, the spectral density of the loss differential at frequency 0. Consequently, we test $H_0 : \mathbb{E}(d_t) = 0,\ \forall\, t$ versus the alternative $H_1 : \mathbb{E}(d_t) \neq 0$. Due to the nestedness of the models, we additionally apply the MSFE F-statistic following McCracken (2007).

Table 6 summarizes our out-of-sample results. The table is subdivided into three panels corresponding to developed, emerging and regional indices. For each panel, we report results for Model 1 and 2. Note that test statistics for Model 2 capture the combined effect from both the consistent as well as opposing signals. In order to be in line with the approach in the in-sample test, we further distinguish between the out-of-sample performance of opposing versus consistent *pe* by reporting separate test results which we label as M 2c for consistent ($D_{i,t}^{oos} = 1$) and M 2o for the opposing signals. By comparing models M 1 and M 2, we find that MAE, RMSE and TIC are quite similar, which suggests at first glance that, if at all, only a small gain in forecast accuracy can be gained by the additional information in Model 2. Out-of-sample R^2, irrespective of the burn-in sample of 6 or 12 years, are positive in all three samples, indicating that returns can be predicted using the *pe* ratio. Whereas for developed countries, Model 2 performs worse than Model 1 based on R^2_{oos}, the information in Model 2 results in better forecast performances for emerging and regional indices, although the forecasts are not significantly different with respect to accuracy, based on the Diebold-Mariano statistic. Contrary to that, using the appropriate test for nested models, the McCracken (2007) F-statistics are above the 95% critical values indicating that we can reject the null of equal predictive accuracy. However, as already noted, the single values of test statistics for Model 2 blurs the effect of distinguishing between consistent and opposing *pe* ratios as it comprises both cases. A more meaningful evaluation of the predictive accuracy is to assess the out-of-sample performance separately for forecasts made from consistent and those made from opposing *pe* ratios. By differentiating between these two cases, we can make a direct comparison between consistent and opposing signal forecasts analogous to the in-sample tests. In line with the results from predictive regression, there are sizeable differences with respect to forecast accuracy for M 2c and M 2o. MAE, RMSE and TIC are considerable lower for predictions using consistent signals, suggesting higher forecast accuracy. These results are similar even across developed, emerging or regional indices. Even more, R^2_{oos} is substantially higher for consistent signal forecasts. For developed countries and regional indices, the forecasts based on opposing signals perform worse than the historical average forecasts (negative $R^2_{6\rightarrow oos}$). Therefore, even out-of-sample results indicate that there are differences in forecastability resulting from the joint signal in the time-series and cross-section.

Taken together, the OOS results support our in-sample test results as well as the preliminary findings from the partial adjustment model. It remains to be shown if the econometric evidence for an improvement in predictive accuracy can also be turned into an economic benefit. This is the focus of the next section.

[Insert Table 6 near here]

6. Investment Performance

6.1. Average utility gain

To value the usefulness of forecasts for economic agents, we first apply a utility-based metric similar to Campbell and Thompson (2008). Cenesizoglu and Timmermann (2012), among others, show that the relationship between a return prediction's out-of-sample statistical performance and the ability to add economic value is rather weak. We therefore employ the measure of an average utility gain for a mean-variance investor with risk aversion γ who allocates the share of a portfolio between stocks and three month US treasury bills. Assume, an investor tries to maximize:

$$U(R_p) = \mathbb{E}(R_p) - \frac{1}{2}\gamma\, \mathbb{V}(R_p).$$

For comparison we compute the certainty equivalent return (CER) for an investor without the forecasts information (historical average return) and for the two forecasts Model 1 (Eq. 14) and Model 2 (Eq. 15).

Each year an investor allocates the following share to equities,

$$a_{i,t} = \left(\frac{1}{\gamma}\right)\left(\frac{\hat{r}_{i,t+1}}{\hat{\sigma}^2_{t+1}}\right),$$

where volatility $\hat{\sigma}^2_{t+1}$ is estimated by a 5-year rolling window of historical returns following Campbell and Thompson (2008). Weights are constrained such that $a_{i,t} \in [0, 1.50]$. Over the out-of-sample period the investor realizes the average utility,

$$\hat{\nu}_i = \hat{\mu}_i - \left(\frac{\gamma}{2}\right)\hat{\sigma}^2_i.$$

The utility gain is then the difference of $\hat{\nu}_i$ to the average utility of the historical average forecast or, put differently, to the constant expected return model.

Table 6 column 7 shows values for the average gain in utility (ΔCER annualized in %), for both models. Discouraging negative values are obtained for all samples and models. Only relatively, ΔCER is higher for Model 2 in all three sample groups. The negative values for ΔCER are somewhat disconcerting. One possible issue might stem from risk attitudes. Altering risk preferences by increasing the parameter γ yields smaller but still negative values for ΔCER. At first glance, this should shed some doubt on whether, on average, predictability can be exploited by mean-variance investors.[22] Thus, we further investigate a

[22] We refrain from testing samples consisting only of consistent or opposing signal predictions (M2c and

real-time global asset allocation exercise as a practical test of our research question.

6.2. Bayesian investment strategy

To further evaluate the predictive content of *pe* ratios we apply an investment strategy using Bayesian updating. The idea is to incorporate return predictions from previous out-of-sample models (Eqs. 14 and 15) into the asset allocation process. Black and Litterman (1991, 1992) provide a seminal approach to blend investors' views with prior (equilibrium) information. The starting point is to specify equilibrium returns through reverse optimization. Consider a standard mean-variance utility function as in the previous section in matrix notation,

$$U = \mathbf{w}'\boldsymbol{\pi} - \left(\frac{\gamma}{2}\right) \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}, \tag{16}$$

where $\mathbf{w}$ is the vector of weights invested in each asset, $\boldsymbol{\pi}$ denotes the vector of equilibrium excess returns and γ is the risk aversion parameter. $\boldsymbol{\Sigma}$ is the covariance matrix of excess returns. Given U is a concave function we can solve for a single global maxima. Taking the first derivative of Eq. (16) with respect to the weights and setting to zero, $\frac{dU}{d\mathbf{w}} = \boldsymbol{\pi} - \gamma\boldsymbol{\Sigma}\mathbf{w} = 0$. Solving for $\boldsymbol{\pi}$ yields

$$\boldsymbol{\pi} = \gamma\, \boldsymbol{\Sigma}\, \mathbf{w}_M. \tag{17}$$

By specifying the risk aversion to γ=2.5 (see e.g. Black and Litterman, 1992)[23] and using the market capitalization of assets as weights $\mathbf{w}_M$, we can determine the CAPM equilibrium returns.[24]

Each period we blend the prior with specific views and uncertainty from predictive regressions. These views $\mathbf{q}$ are either relative or absolute. The relative view asserts that lower *pe* ratios outperform higher *pe* ratios by the average estimate of the predictive regression ($\hat{r}$). The absolute view is specified in the way that for each asset (country), we use the estimate $\hat{r}_i$ in the vector $\mathbf{q}$.

The confidence in the views is inversely specified in the diagonal matrix $\boldsymbol{\Omega}$. We proxy our confidence in the return forecasts (views $\mathbf{q}$) proportional to the scaled R^2 of the predictive regression.[25]

M2o) mainly because of different sample sizes and potential biases in test statistics (DM-stat, MSFE F-stat, and CER).

[23] Again, altering the risk aversion parameter γ between 1 and 10 does not change our results qualitatively.

[24] We use the market capitalization of a developed market index (MSCI World) as a proxy for the market portfolio. Similarly, for developing countries we use the MSCI Emerging Markets index as a proxy. Both indices are gathered from Thomson Reuters Datastream to be consistent with previous data.

[25] This approach is similar to Beach and Orlov (2007) who use the variance of the residuals from GARCH factor models to derive confidence estimates.

Given the views and the respective confidence we can formulate the conditional distribution of returns $\boldsymbol{\mu}$ by applying Bayes theorem,[26]

$$\boldsymbol{\mu}|\mathbf{q};\boldsymbol{\Omega} \sim \mathcal{N}(\boldsymbol{\mu}_{BL}, \boldsymbol{\Sigma}_{BL}), \tag{18}$$

where $\boldsymbol{\mu}_{BL}$ is the combined return vector,[27]

$$\boldsymbol{\mu}_{BL} = [(\tau\boldsymbol{\Sigma})^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}]^{-1}[(\tau\boldsymbol{\Sigma})^{-1}\boldsymbol{\pi} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{q}], \tag{21}$$

and $\boldsymbol{\Sigma}_{BL}$ the combined covariance,

$$\boldsymbol{\Sigma}^{\boldsymbol{\mu}}_{BL} = [(\tau\boldsymbol{\Sigma})^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}]^{-1}. \tag{22}$$

τ represents a scaling factor corresponding to the confidence level of the CAPM prior mean which we set to 0.05 following He and Litterman (2002).[28] The next step is to feed the posterior estimates of mean returns and variance into a portfolio optimizer which yields the optimized weights for the portfolio composition,

$$\mathbf{w}^* = \gamma\, \boldsymbol{\Sigma}^{-1}\, \boldsymbol{\mu}_{BL}. \tag{23}$$

We specify a 10 year in-sample pre-estimation period to compute the equilibrium returns and covariances. Then, we rebalance each year the portfolio weights according to the updates from the Black-Litterman model in an expanding window approach. Critically, to mimic an actual investment strategy, we do only incorporate prior information and estimates from out-of-sample forecasts. A constraint is imposed such that weights sum to one.

The results of the global asset allocation exercise are shown in Figures 7 and 8 for portfolios comprising developed and emerging country indices respectively. Figure 7 depicts the out-of-sample performance (cumulative returns) of portfolios starting in 1983. All but one portfolio outperform the benchmarks index, the MSCI World. Clearly, the Black-Litterman optimized portfolios using absolute views perform better in terms of cumulative returns. The equilibrium portfolio composed by market capitalization weights perform slightly better

[26]For the derivation see e.g. Satchell and Scowcroft (2000) or Meucci (2010).

[27]For computational purposes we follow Meucci (2010) who uses a more stable representation of the posterior parameters,

$$\boldsymbol{\mu}_{BL} = \boldsymbol{\pi} + \tau\boldsymbol{\Sigma}\mathbf{P}'(\tau\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}' + \boldsymbol{\Omega})^{-1}(\mathbf{q} - \mathbf{P}\boldsymbol{\pi}) \tag{19}$$

$$\boldsymbol{\Sigma}^{\boldsymbol{\mu}}_{BL} = (1+\tau)\boldsymbol{\Sigma} - \tau^2\boldsymbol{\Sigma}\mathbf{P}'(\tau\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}' + \boldsymbol{\Omega})^{-1}\mathbf{P}\boldsymbol{\Sigma}. \tag{20}$$

[28]Some authors use different values for τ such as $\tau = \frac{1}{T}$ (Meucci, 2010). Again, our results are robust against variations of this parameter.

than the benchmark index, which might be a result of slightly different composition of the portfolios. Model 2, both with absolute views and relative views, performs better indicating a small but existing performance gain from the information about consistent and opposing signals. Figure 8 highlights the same exercise for developing country index portfolios starting in 2001. Again, absolute views dominate others in terms of cumulative returns. However, Model 1 with absolute views clearly outperforms other portfolios.

To better evaluate the risk return relationship of the portfolio, we determined the most common performance statistics in Table 7. For developed countries, Model 2 with absolute views dominates others based on mean return, Sharpe ratio, Information ratio, Jensen's alpha, the maximum drawdown and the lower partial moment (Sortino ratio). The performance gains are economically meanigful with a Sharpe ratio of 38.5% compared to 28% of the benchmark index. All but Model 2 relative exhibit positive information ratios with respect to the benchmark. For emerging portfolios, Model 1 absolute performs best based on all listed statistics. The Sharpe ratio increases from 31.9% to 60.2%., i.e. almost doubles compared to the benchmark.

The global asset allocation exercise with Bayesian updating demonstrates that the econometric evidence of predictability can actually be tunred into an even stronger economic significance in terms of investment performance measures. This result is in line with e.g. Campbell and Thompson (2008), who already noted a magnification effect which can turn a small improvement in predictive accuracy into a significantly larger utility gain for an investor. Interestingly, our results for Model 2, which uses the consistent/opposing dummy, can not be taken as evidence that the explicit exploitation of consistent signals leads to an improvement in the investment performance. At first sight, this seems to be at odds with our results in the previous sections. However, recall that by construction of the Bayesian updating process, the impact of predicted returns in the asset allocation depends on the confidence we put on these forecasts. As noted above, we proxy our confidence in the prediction by their proportional R^2 in the out-of-sample predictive regression. Countries with R^2s close to zero will therefore not obtain a different weighting as in the benchmark. Thus, the performance difference to the benchmark will only be driven by countries with non-zero R^2. As we have shown in our out-of-sample tests, high OOS R^2 are only obtained for countries which display a consistent *pe* ratio. Therefore, although Model 1 is an unconditional model, it captures already the effect that if an improvement in the investment performance can be obtained, then it results from the prediction on the basis of consistent *pe* ratios. For this reason, Model 2 can by construction of the Black-Litterman approach not improve significantly over Model 1. To emphasize this point, we run Model 2 but shut down views whenever countries obtain opposing signals for their previous *pe* ratios. This is achieved by setting the confidence in views for opposing signals to virtually zero, resulting in market weights. The results of this exercise are shown in Table 7 column 5, M2c abs. The performance measures change only insignificantly compared to Model 1 and 2 with absolute views, supporting our conjecture that opposing signals do not strengthen portfolio performance. Taken together, our results from the Bayesian investment strategy support our econometrical findings from the previous sections.

[Insert Table 7 near here]

[Insert Figure 7 and 8 near here]

7. Further Robustness

In this section we provide additional robustness checks. We extend the previous analysis to the dividend yield (*dy*) and a price to cash flow ratio (*pc*). Additionally, subsample results are provided for out-of-sample measures.

Data for the dividend yield (*dy*) and the price to cash flow ratio (*pc*) is gathered from Datastream (data type code "DY" and "PC"). *pc* is defined as the cash earnings per share adjusted for capital changes. Cash flow is calculated as funds from operations which is the sum of net income and all non-cash charges or credits according to Datastream. PC ratios are available only from 1983M01 onwards for developed countries. Again we take logs of the ratios emphasized by lower case letters.

Table 8 shows results of the partial adjustment model using *dy* and *pc*. In line with the results for the *pe* ratios the results suggest no significant partial adjustment for opposing signal ratios (column 3). Ratios with consistent signals display strong and significant partial adjustment to the mean, supporting the results of the previous analysis. Some differences can be detected for *dy* with respect to significance, particular for emerging countries (columns 5-8). Results for *pc* (Panel D) are very similar to results for *pe*, suggesting a robust difference when we condition on consistent versus opposing signals.

[Insert Table 8 near here]

Table 9 displays results for the pooled predictive regressions. Both for *dy* and *pc*, opposing signal predictions are not significant and display a very low R^2. Again, comparable to *pe* ratios, strong predictability can be detected for consistent signal *dy* and *pc* ratios across country groups. A particularly high R^2 of 1.2% is displayed by the *pc* ratio as predictor at the 1 month horizon suggesting a potential superior predictive relation of *pc* ratios compared to the other predictor variables.

[Insert Table 9 near here]

Table 10 shows out-of-sample statistics for *dy* and *pc*. MAE, RMSE and TIC are comparable to the results for *pe* ratios. Model 2 including the consistent signal does not improve forecast accuracy. But again, as we decompose the results for opposing and consistent signal forecasts, the latter display more accuracy. R^2_{oos} are positive for both models and all country groups/ratios. In line with previous evidence we find that opposing signal forecasts cannot beat the historical average forecasts as emphasized by M2o. DM *p*-values suggest no statistical difference between Model 1 and Model 2. The F-statistics for emerging countries and regions, however, are higher than the 95% critical value from table 4 in McCracken (2007) of 1.706 and thus we can reject the null of equal predictive accuracy for these samples. Certainty equivalent gains are positive for Model 1 *dy* developed and for the *pc* ratio. Other samples display negative CER gains just as the results for *pe* ratios. The last three columns highlight differences arising from variation in the risk aversion coefficient. Variations are only in magnitude but no signs change.

[Insert Table 10 near here]

Following Goyal and Welch (2008), we additionally employ graphs for the cumulative squared prediction error of the null hypothesis (benchmark mean forecast) minus the cumulative squared prediction error of Model M2c (consistent signals). Figure 9 shows results for *pe*, *dy* and *pc*. We choose an out-of-sample period from 1993M01 to 2014M04 in order to compare directly the three predictor variables. The solid line marks the mean of cumulative differences across developed countries. All three predictor variables beat the historical average forecast in this exercise. However, during the recent financial crisis the models perform worse than the benchmark forecast and previous cumulative gains are almost wiped out. Interestingly, the *pc* ratio displays a particularly good performance across the OOS horizon.

[Insert Figure 9 near here]

Table 11 shows sub-sample results for all three predictor variables. Columns 2 and 3 highlight in-sample measures across periods. The regression coefficient $\beta^c_{k=1m}$ denotes the 1 month ahead forecast for consistent ratios. Estimates are highly significant for most periods. Interestingly, during the sub-period from 2005 to 2014 the coefficient for *pe* and *dy* is insignificant and, even worse, displays the wrong sign. We suspect the financial crisis within this period disrupts our estimates. R^2 mirrors these results. The same pattern is also found in out-of-sample measures. The last two columns display Sharpe ratios and Jensen's alpha. Largest alphas are generated in the early sample 1983 to 1993. Again, during 2005-2014, alphas are below 1% for *pe* and *dy*. Overall, the *pc* ratio seems to perform best compared to the other two ratios, displaying an overall alpha of 3.99% for 1994-2014, compared to 2.59% and 1.46% for *pe* and *dy* respectively.

[**Insert Table 11 near here**]

8. Conclusion

Inspired by practical investment advice in international asset allocation, we ask if return predictions on the basis of aggregate price to fundamental ratios can be improved by the combined time-series and cross-section information. Essentially, we address the question if the cross-section offers some additional predictive accuracy. To answer this question, we classify current *pe* ratios as consistent if they are both above (below) the long-run domestic mean as well as above (below) the global mean and as opposing otherwise. On the basis of this sample split, we analyze *pe* ratios in four directions: First, as preliminary evidence we estimate partial adjustment models to see if mean-reverting behavior is different. Second, in-sample predictive regressions are run. Third, we offer out-of-sample tests, and fourth we assess the economic impact of predictability in a Bayesian asset allocation strategy. In each of the four sections, we find clear differences between consistent and opposing *pe* ratios.
We show that only consistent *pe* ratios display a significant mean-reverting behavior. By calibrating a simulation exercise, we rule out that the results are mechanical. As a second and major contribution, we run predictive regressions. Contrary to the majority of previous work, we run panel regressions with an error term modeled as AR(1) and we use standard errors from Seemingly Unrelated Regression (SUR), which ensures that our results are conservative estimates. We find that for opposing *pe* ratios none of the specifications yield a significant coefficient, while for consistent *pe* ratios all specifications are strongly significant and show higher R^2 than in the unconditional setting. In general, positive in-sample results are certainly not a guarantee for good out-of-sample results. It is known that in-sample tests have higher power in many situations and that out-of-sample results are often poor because there is too little information in small samples to improve the return forecast (see e.g. Inoue and Kilian 2005, Campbell and Thompson 2008 or Lettau and Ludvigson 2009). Against this background, our out-of-sample results are surprisingly good. Compared to the in-sample evidence, out-of-sample results are weaker but even in the unconditional model,

we find small but positive OOS R^2s. The interesting finding however is the separate evaluation of the OOS performance of consistent versus opposing *pe* ratios. It turns out that the sample of consistent *pe* ratios is able to improve all OOS test statistics in a sizeable way, confirming the superior accuracy from adding cross-sectional information. Finally, feeding the forecasts into an asset allocation strategy along the lines of the Black and Litterman (1991, 1992) model confirms that the predictability evidence can be exploited by a real-time investor to realize a significant economic benefit in terms of outperformance against a benchmark. These results are robust against using different predictor variables such as *dy* and *pc* which exhibit similar patterns. Even more, results are stronger for *pc* ratios, suggesting that our approach is not at all limited to the *pe* ratio.

Our results provide the following general implications. First and foremost, we show that the cross-sectional information provides some additional predictive power and is able to improve the predictive accuracy in return forecasts. Second, we observe that a specific country's price to fundamental ratio will switch from being consistent to opposing over time. By combining our finding that only price to fundamental ratios that can be classified as consistent are able to provide significant predictive power leads to the implication that the predictability within individual countries will be unstable over time. Thus, our results offer a new approach to identify time periods in which return predictions are expected to work. Related to this, our results can also explain the heterogeneous results for predictability in the cross-section, since only countries with consistent price to fundamental ratios are found to provide significant predictive coefficients. A full exploration of temporal instability and cross-sectional heterogeneity is left for future research.

Appendix A. Simulation Exercise

This section details the simulation approach. We generate artificial *pe* ratios using an Ornstein-Uhlenbeck process.

In order to simulate *pe* ratios close to the empirical ones, we match the model parameters to the empirical estimates. This leads to an extended Ornstein-Uhlenbeck model with starting value,

$$pe_0 = \mu + \sigma\, dW_t,$$

and the stochastic differential equation

$$d(pe)_t = \theta(\mu - pe_{t-1})\, dt + \sigma\, dW_t,$$

where parameters θ and μ are random variables themselves matched with empirical estimates.

Mean empirical estimates from pe ratios are $\widehat{\theta} = 0.98$, $\widehat{\mu} = 2.96$ and $\widehat{\sigma} = 0.42$.

θ, the rate by which shocks dissipate, emphasizes the high persistence in empirical data and therefore we model it close to unit root properties. μ indicates the diverse means for each artificial pe ratios. The rationale to use additional randomness in parameters stems from similar observed phenomena in empirical pe ratios. With the resulting extended Ornstein-Uhlenbeck process we generate ratios with very similar (higher) moments to empirical ratios.

We simulate 15 artificial pe ratios for the same horizon as the empirical ones. Table 12 shows results for both the partial adjustment parameter $\widehat{\lambda}$ and the predictive coefficient $\widehat{\beta}$. Partial adjustment speeds for consistent time-series and cross-sectional signals differ only slightly to opposing signals. Both exhibit positive and highly significant adjustment speeds. When it comes to return predictability with simulated pe ratios, we find no predictive power. Predictive coefficients $\widehat{\beta}$ are not significant for neither consistent nor opposing signals. Even more, R^2 is virtually zero for both regressions. This shows that some persistence in predictive variables alone does not explain the return predictability evidence. Importantly, this simulation exercise shows that the findings with empirical data are not mechanical. Furthermore, simulations with constant parameters θ and μ yield very similar, almost identical, results.

[Insert Table 12 near here]

Appendix B. Consistent versus Opposing Signals

Figure 10 shows the distribution of opposing versus consistent signals for each developed country's pe ratios from 1973 to 2014. The black bars mark opposing signals and the grey bars mark consistent signals. The graph emphasizes that the two conditions change frequently for each country across time. Long periods of consistent signals are detected for Japan which has for many years been overvalued both in the time-series and the cross-sectional dimension. Long periods of opposing signals exhibit the Netherlands and Singapore. All in all, of the 7380 monthly observations around 66% of pe ratios are consistent whereas 34% are opposing.

[Insert Figure 10 near here]

References

Ahn, D.-H., Boudoukh, J., Richardson, M., Whitelaw, R. F., 2002. Partial adjustment or stale prices? Implications from stock index and futures return autocorrelations. Review of Financial Studies 15, pp.655–689.

Amihud, Y., Hurvich, C. M., 2004. Predictive regressions: A reduced-bias estimation method. The Journal of Financial and Quantitative Analysis 39, pp.813–841.

Ang, A., Bekaert, G., 2007. Stock return predictability: Is it there? Review of Financial Studies 20, pp.651–707.

Ang, A., Liu, J., 2007. Risk, return, and dividends. Journal of Financial Economics 85, pp.1–38.

Beach, S., Orlov, A., 2007. An application of the Black-Litterman model with EGARCH-M-derived views for international portfolio management. Financial Markets and Portfolio Management 21, pp.147–166.

Beck, N., Katz, J. N., 1995. What to do (and not to do) with time-series cross-section data. The American Political Science Review 89, pp.634–647.

Bekaert, G., Hodrick, R. J., 1992. Characterizing predictable components in excess returns on equity and foreign exchange markets. The Journal of Finance 47, pp.467–509.

Black, F., Litterman, R., 1992. Global portfolio optimization. Financial Analysts Journal 48, pp.28–43.

Black, F., Litterman, R. B., 1991. Asset allocation: Combining investor views with market equilibrium. Journal of Fixed Income 1, pp.7–18.

Bollerslev, T., Marrone, J., Xu, L., Zhou, H., 2014. Stock return predictability and variance risk premia: Statistical inference and international evidence. Journal of Financial and Quantitative Analysis 49, pp.633–661.

Boudoukh, J., Michaely, R., Richardson, M., Roberts, M. R., 2007. On the importance of measuring payout yield: Implications for empirical asset pricing. The Journal of Finance 62, pp.877–915.

Boudoukh, J., Richardson, M., Whitelaw, R. F., 2008. The myth of long-horizon predictability. Review of Financial Studies 21, pp.1577–1605.

Campbell, J. Y., Hamao, Y., 1992. Predictable stock returns in the United States and Japan: A study of long-term capital market integration. The Journal of Finance 47, pp.43–69.

Campbell, J. Y., Shiller, R. J., 1988a. The dividend-price ratio and expectations of future dividends and discount factors. Review of Financial Studies 1, pp.195–228.

Campbell, J. Y., Shiller, R. J., 1988b. Stock prices, earnings, and expected dividends. The Journal of Finance 43, pp.661–676.

Campbell, J. Y., Shiller, R. J., 2001. Valuation ratios and the long-run stock market outlook: An update. Journal of Portfolio Management 24, pp.11–26.

Campbell, J. Y., Thompson, S. B., 2008. Predicting excess stock returns out of sample: Can anything beat the historical average? Review of Financial Studies 21, pp.1509–1531.

Campbell, J. Y., Yogo, M., 2006. Efficient tests of stock return predictability. Journal of Financial Economics 81, pp.27–60.

Cenesizoglu, T., Timmermann, A., 2012. Do return prediction models add economic value? Journal of Banking & Finance 36, pp.2974–2987.

Chen, L., Da, Z., Priestley, R., 2012. Dividend smoothing and predictability. Management Science 58, pp.1834–1853.

Cochrane, J. H., 1999. New facts in finance. Economic Perspectives pp. 36–58.

Cochrane, J. H., 2005. Asset Pricing. Princeton University Press, revised ed.

Cochrane, J. H., 2008. The dog that did not bark: A defense of return predictability. Review of Financial Studies 21, pp.1533–1575.

Cochrane, J. H., 2011. Presidential address: Discount rates. The Journal of Finance 66, pp.1047–1108.

Cremers, K. M., 2002. Stock return predictability: A bayesian model selection perspective. Review of Financial Studies 15, pp.1223–1249.

Cutler, D. M., Poterba, J. M., Summers, L. H., 1991. Speculative dynamics. The Review of Economic Studies 58, pp.529–546.

Dangl, T., Halling, M., 2012. Predictive regressions with time-varying coefficients. Journal of Financial Economics 106, pp.157–181.

Diebold, F. X., Mariano, R. S., 1995. Comparing predictive accuracy. Journal of Business & Economic Statistics 13, pp.253–263.

Engsted, T., Pedersen, T. Q., 2010. The dividend-price ratio does predict dividend growth: International evidence. Journal of Empirical Finance 17, pp.585–605.

Fama, E. F., French, K. R., 1988. Dividend yields and expected stock returns. Journal of Financial Economics 22, pp.3–25.

Fama, E. F., French, K. R., 2000. Forecasting profitability and earnings. The Journal of Business 73, pp.161–175.

Fama, E. F., French, K. R., 2001. Disappearing dividends: Changing firm characteristics or lower propensity to pay? Journal of Applied Corporate Finance 14, pp.67–79.

Ferreira, M. A., Santa-Clara, P., 2011. Forecasting stock market returns: The sum of the parts is more than the whole. Journal of Financial Economics 100, pp.514–537.

Ferson, W. E., Harvey, C. R., 1993. The risk and predictability of international equity returns. Review of Financial Studies 6, pp.527–566.

Flannery, M. J., Rangan, K. P., 2006. Partial adjustment toward target capital structures. Journal of Financial Economics 79, pp.469–506.

Giot, P., Petitjean, M., 2006. International stock return predictability: Statistical evidence and economic significance. CORE discussion paper: Center for Operations Research and Econometrics, CORE.

Giot, P., Petitjean, M., 2011. On the statistical and economic performance of stock return predictive regression models: an international perspective. Quantitative Finance 11, 175–193.

Goyal, A., Welch, I., 2008. A comprehensive look at the empirical performance of equity premium prediction. Review of Financial Studies 21, pp.1455–1508.

Graham, B., Dodd, D. L., 1934. Security Analysis. Whittlesey House, first ed.

Han, Y., Yang, K., Zhou, G., 2013. A new anomaly: The cross-sectional profitability of technical analysis. Journal of Financial and Quantitative Analysis 48, pp.1433–1461.

Harvey, C., 1991. The world price of covariance risk. The Journal of Finance 46, pp.111–157.

Harvey, C., 1995. Predictable risk and returns in emerging markets. Review of Financial Studies 8, pp.773–816.

He, G., Litterman, B., 2002. The intuition behind Black-Litterman model portfolios. SSRN 8221.

Henkel, S. J., Martin, J. S., Nardari, F., 2011. Time-varying short-horizon predictability. Journal of Financial Economics 99, pp.560–580.

Hjalmarsson, E., 2010. Predicting global stock returns. Journal of Financial and Quantitative Analysis 45, pp.49–80.

Hodrick, R., 1992. Dividend yields and expected stock returns: Alternative procedures for inference and measurement. Review of Financial Studies 5, pp.357–386.

Im, K. S., Pesaran, M., Shin, Y., 2003. Testing for unit roots in heterogeneous panels. Journal of Econometrics 115, pp.53–74.

Inoue, A., Kilian, L., 2005. In-sample or out-of-sample tests of predictability: Which one should we use? Econometric Reviews 23, pp.371–402.

Jordan, S. J., Vivian, A., Wohar, M. E., 2015. Location, location, location: Currency effects and return predictability? Applied Economics 47, pp.1883–1898.

Jorion, P., Goetzmann, W. N., 1999. Global stock markets in the twentieth century. The Journal of Finance 54, pp.953–980.

Kellard, N. M., Nankervis, J. C., Papadimitriou, F. I., 2010. Predicting the equity premium with dividend ratios: Reconciling the evidence. Journal of Empirical Finance 17, pp.539–551.

Koijen, R. S., Van Nieuwerburgh, S., 2011. Predictability of returns and cash flows. Annual Review of Financial Economics 3, pp.467–491.

Koustas, Z., Serletis, A., 2005. Rational bubbles or persistent deviations from market fundamentals? Journal of Banking & Finance 29, pp.2523–2539.

Kwiatkowski, D., Phillips, P. C., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? Journal of Econometrics 54, pp.159–178.

Lamont, O., 1998. Earnings and expected returns. The Journal of Finance 53, pp.1563–1587.

Lettau, M., Ludvigson, S. C., 2009. Measuring and Modeling Variation in the Risk-Return Trade-off. In: Ait-Sahalia, Y., Hansen, L. (eds.), *Handbook of Financial Econometrics, Vol. 1: Tools and Techniques (Handbooks in Finance)*, North Holland, pp. 617–690.

Lettau, M., Van Nieuwerburgh, S., 2008. Reconciling the return predictability evidence. Review of Financial Studies 21, pp.1607–1652.

Lev, B., 1969. Industry averages as targets for financial ratios. Journal of Accounting Research 7, pp.290–299.

Levin, A., Lin, C.-F., Chu, C.-S. J., 2002. Unit root tests in panel data: Asymptotic and finite-sample properties. Journal of Econometrics 108, pp.1–24.

Lewellen, J., 2004. Predicting returns with financial ratios. Journal of Financial Economics 74, pp.209–235.

McCracken, M. W., 2007. Asymptotics for out of sample tests of granger causality. Journal of Econometrics 140, pp.719–752.

Mcmillan, D. G., Wohar, M. E., 2013. A Panel Analysis Of The Stock ReturnDividend Yield Relation: Predicting Returns And Dividend Growth. Manchester School 81, pp.386–400.

Meucci, A., 2010. Black-Litterman Approach, in Encyclopedia of Quantitative Finance, John Wiley & Sons, Ltd.

Nagayasu, J., 2007. Putting the dividendprice ratio under the microscope. Finance Research Letters 4, pp.186–195.

Neely, C. J., Rapach, D. E., Tu, J., Zhou, G., 2014. Forecasting the equity risk premium: The role of technical indicators. Management Science 60, pp.1772–1791.

Newey, W. K., West, K. D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, pp.703–708.

Park, C., 2010. When does the dividendprice ratio predict stock returns? Journal of Empirical Finance 17, pp.81–101.

Paye, B. S., Timmermann, A., 2006. Instability of return prediction models. Journal of Empirical Finance 13, pp.274–315.

Pettenuzzo, D., Timmermann, A., Valkanov, R., 2014. Forecasting stock returns under economic constraints. Journal of Financial Economics 114, pp.517–553.

Phillips, P. C. B., Perron, P., 1988. Testing for a unit root in time series regression. Biometrika 75, pp.335–346.

Poterba, J. M., Summers, L. H., 1988. Mean reversion in stock prices: Evidence and implications. Journal of Financial Economics 22, pp.27–59.

Rapach, D. E., Strauss, J. K., Zhou, G., 2010. Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. Review of Financial Studies 23, pp.821–862.

Rapach, D. E., Strauss, J. K., Zhou, G., 2013. International stock return predictability: What is the role of the United States? The Journal of Finance 68, pp.1633–1662.

Rapach, D. E., Wohar, M. E., Rangvid, J., 2005. Macro variables and international stock return predictability. International Journal of Forecasting 21, pp.137–166.

Rapach, D. E., Zhou, G., 2013. Chapter 6 - Forecasting stock returns. In: Elliott, G., Timmermann, A. (eds.), *Handbook of Economic Forecasting*, Elsevier, vol. 2, Part A of *Handbook of Economic Forecasting*, pp. 328–383.

Richards, A. J., 1995. Comovements in national stock market returns: Evidence of predictability, but not cointegration. Journal of Monetary Economics 36, pp.631–654.

Rozeff, M. S., 1984. Dividend yields are equity risk premiums. The Journal of Portfolio Management 11, pp.68–75.

Said, S. E., Dickey, D. A., 1984. Testing for unit roots in autoregressive-moving average models of unknown order. Biometrika 71, pp.599–607.

Satchell, S., Scowcroft, A., 2000. A demystification of the BlackLitterman model: Managing quantitative and traditional portfolio construction. Journal of Asset Management 1, pp.138–150.

Schrimpf, A., 2010. International stock return predictability under model uncertainty. Journal of International Money and Finance 29, pp.1256–1282.

Solnik, B., 1993. The performance of international asset allocation strategies using conditioning information. Journal of Empirical Finance 1, pp.33–55.

Stambaugh, R. F., 1999. Predictive regressions. Journal of Financial Economics 54, pp.375–421.

Turner, J. A., 2015. Casting doubt on the predictability of stock returns in real time: Bayesian model averaging using realistic priors. Review of Finance 19, pp.785–821.

Valkanov, R., 2003. Long-horizon regressions: Theoretical results and applications. Journal of Financial Economics 68, pp.201–232.

Van Binsbergen, J. H., Koijen, R. S. J., 2010. Predictive regressions: A present-value approach. The Journal of Finance 65, pp.1439–1471.

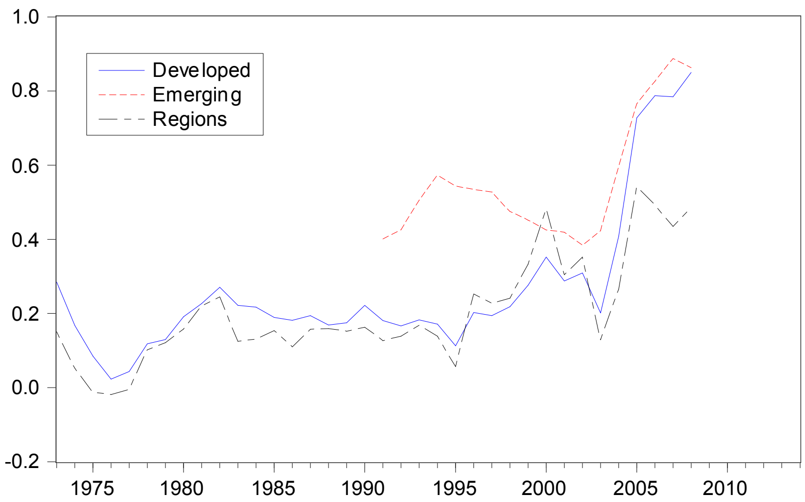

Fig. 2. Partial adjustment speeds of pe ratios to the group mean $\overline{pe}$

This figure shows partial adjustment of pe ratios to the group mean $\overline{pe}$. Coefficient $\widehat{\lambda_k}$ of the model $pe_{i,t+k} = \overline{pe}\lambda_k + (1 - \lambda_k)\, pe_{i,t} + \varepsilon_{i,t+k}$. Rolling window of 5 years, 1 year steps.

Table 1: Summary statistics

This table shows summary statistics for *pe* ratios and monthly returns from the total return index. Mean returns and standard deviations are in %.

	pe Ratios						Returns			
	Mean	Max	Min	Std.D.	Skew	Kurt	Mean	Std.D.	Skew	Kurt
Panel A: Developed Countries 1973M01-2014M05										
Australia	14.83	24.70	5.60	4.56	0.13	2.11	0.86	7.33	-1.06	8.06
Austria	16.98	32.60	5.70	4.65	0.90	4.24	0.79	6.69	-0.30	8.00
Belgium	13.34	37.20	4.90	4.38	1.62	8.11	0.89	5.96	-0.77	7.34
Canada	15.54	31.70	6.10	4.79	0.10	2.81	0.80	5.62	-0.86	6.51
Denmark	16.96	60.30	5.40	7.83	1.87	9.37	1.02	5.99	-0.55	4.89
France	13.41	28.20	5.70	4.23	0.67	3.24	0.96	6.82	-0.59	4.67
Germany	15.13	26.80	8.40	3.69	0.50	2.62	0.84	6.08	-0.51	4.40
Hongkong	14.42	73.40	5.50	5.55	4.02	34.96	0.96	6.82	-0.58	9.87
Ireland	12.42	26.20	3.90	4.92	0.25	2.20	0.90	7.15	-0.34	5.75
Japan	35.39	84.60	13.00	17.41	0.68	2.38	0.60	6.21	0.03	4.08
Netherland	12.92	37.30	4.20	5.52	1.11	4.45	0.95	5.79	-1.14	8.18
Singapore	18.85	54.60	5.00	6.64	1.54	8.10	0.72	8.30	-0.17	7.24
Switzerland	14.56	28.80	6.00	4.51	0.48	2.64	0.96	5.79	-0.51	4.35
UK	13.65	26.70	3.20	4.07	0.47	3.39	0.90	6.21	0.27	7.92
US	16.31	31.60	6.80	5.75	0.44	2.70	0.84	4.63	-0.75	5.99
Mean	16.31						0.87			
Panel B: Emerging/Developing Countries 1991M02-2014M05										
Chile	17.66	26.00	8.00	3.15	-0.19	3.41	1.07	6.87	-0.33	4.33
India	18.99	61.10	8.60	7.59	1.96	8.48	0.76	10.44	-0.01	4.86
Greece	15.52	36.30	2.00	6.71	0.88	3.45	0.18	9.69	-0.35	4.55
Indonesia	16.41	33.10	7.10	4.75	0.24	3.23	0.29	17.44	-0.40	10.56
Korea	15.28	38.40	6.50	5.67	1.37	5.17	0.56	10.57	0.03	4.90
Malaysia	17.51	42.90	7.00	5.22	1.20	5.08	0.69	8.28	-0.45	8.84
Mexico	13.68	21.30	6.70	3.00	0.35	3.22	1.10	8.75	-1.16	6.76
Taiwan	20.17	37.50	8.00	6.57	0.58	2.74	0.43	8.73	0.49	4.91
Thailand	14.34	55.10	5.60	6.91	2.70	13.09	0.68	10.48	-0.10	5.59
Sri Lanka	13.42	33.20	3.90	5.72	0.57	2.98	0.71	7.87	0.10	4.21
South Africa	14.89	23.90	8.30	3.29	0.27	2.88	0.82	7.96	-0.87	7.00
Philippines	17.52	37.00	8.30	5.37	1.17	4.81	0.91	8.23	0.15	6.31
Mean	16.28						0.68			
Panel C: Regions 1973M01-2014M05										
Americas	15.81	30.40	6.80	5.25	0.42	2.81	0.83	4.67	-0.79	6.16
Asia	27.81	67.90	10.50	12.27	1.00	3.27	0.71	5.94	-0.08	4.01
Australasia	14.61	24.30	5.50	4.43	0.14	2.12	0.87	7.28	-1.07	8.19
Developed	17.45	31.70	8.00	5.40	0.18	2.27	0.80	4.56	-0.62	4.80
EMU	13.89	26.40	6.90	3.68	0.66	3.47	0.87	5.60	-0.69	4.94
Europe	13.52	25.80	5.80	3.84	0.56	3.22	0.90	5.39	-0.64	5.65
Far East	29.85	69.30	10.90	12.82	0.96	2.94	0.67	5.90	0.01	4.26
North America	16.25	31.20	6.80	5.63	0.41	2.65	0.83	4.46	-0.77	6.12
NORCS	14.12	29.50	6.60	4.42	0.77	3.76	0.84	5.74	-0.95	6.85
Pacific Basin	26.72	63.30	10.40	11.49	0.95	3.11	0.72	5.85	-0.14	4.04
PIIGS	13.02	39.50	3.40	5.83	0.52	3.53	0.77	7.11	-0.32	4.28
Scandinavia	14.96	58.30	5.80	7.18	2.88	14.14	0.92	6.55	-0.57	4.66
SE Asia	18.32	54.30	6.50	6.08	1.98	10.43	0.72	8.37	-0.24	7.14
World	16.87	30.40	8.00	5.01	0.20	2.39	0.80	4.59	-0.66	5.11
Mean	18.09						0.80			

Table 2: Unit root / stationarity tests of log (P/E)

This table shows the autocorrelation coefficient for the first lag (AC(1)) and the following unit root tests. ADF is the Augmented Dickey Fuller test (Said and Dickey, 1984) with H_0 =unit root. The KPSS test (Kwiatkowski et al., 1992) tests H_0 =stationary and the Phillips Perron statistic tests H_0 =unit root (Phillips and Perron, 1988). Panel unit root tests of Levin et al. (2002) and Im et al. (2003) test H_0 =unit root. *,** and *** denote significance at the 10%, 5% and 1% level respectively.

	AC(1)	ADF	KPSS	Phillips Perron
Panel A: Developed Countries 1973M01-2014M05				
Australia	0.981	-2.223	1.640***	-2.319
Austria	0.948	-4.331***	0.345	-4.358***
Belgium	0.957	-4.398***	0.359*	-4.140***
Canada	0.980	-2.731*	1.372***	-2.947*
Denmark	0.947	-3.861***	0.719**	-3.751***
France	0.966	-3.071**	0.977***	-3.083**
Germany	0.954	-3.301**	0.416*	-3.541***
Hongkong	0.935	-4.522***	0.138	-4.665***
Ireland	0.970	-2.277	1.327***	-2.777*
Japan	0.988	-1.679	0.650**	-1.472
Netherland	0.980	-2.620*	1.520***	-2.450
Singapore	0.951	-4.015***	1.462***	-4.273***
Switzerland	0.984	-2.187	1.994***	-2.009
UK	0.978	-2.630*	1.067***	-2.994**
US	0.989	-1.454	1.574***	-1.916
Panel:	Levin, Lin & Chu t*: -0.749 (p-val=0.2269)			
	Im, Pesaran and Shin W-stat: -7.048 (p-val=0.0000)			
Panel B: Emerging/Developing Countries 1991M02-2014M05				
Chile	0.876	-5.214***	0.289	-5.314***
India	0.950	-2.685*	0.480**	-2.840*
Greece	0.932	-1.771	0.285	-2.005
Indonesia	0.928	-3.215**	0.348*	-3.242**
Korea	0.919	-3.891***	0.503**	-3.787***
Malaysia	0.953	-2.596*	0.860***	-2.879**
Mexico	0.903	-3.374**	1.381***	-3.413**
Taiwan	0.945	-3.525***	1.057***	-3.362**
Thailand	0.909	-3.672***	0.393*	-3.785***
Sri Lanka	0.980	-1.684	0.300	-2.046
South Africa	0.951	-2.647*	0.292	-2.694*
Philippines	0.941	-3.675***	0.176	-3.569***
Panel:	Levin, Lin & Chu t*: -1.465 (p-val=0.072)			
	Im, Pesaran and Shin W-stat: -6.578 (p-val=0.000)			
Panel C: Regions 1973M01-2014M05				
Americas	0.988	-1.695	1.551***	-2.037
Asia	0.988	-1.531	0.678**	-1.635
Australasia	0.979	-2.250	1.685***	-2.308
Developed	0.990	-1.969	1.012***	-2.096
EMU	0.973	-3.035**	0.932***	-2.926**
Europe	0.981	-2.593*	1.166***	-2.671*
Far East	0.988	-1.737	0.581**	-1.837
North America	0.989	-1.770	1.574***	-1.944
NORCS	0.974	-2.991**	0.862***	-3.285**
Pacific Basin	0.989	-1.541	0.643**	-1.641
PIIGS	0.979	-2.628*	1.268***	-2.480
Scandinavia	0.962	-3.510***	0.370*	-3.840***
SE Asia	0.942	-5.059***	1.478***	-4.546***
World	0.989	-2.090	0.961***	-2.162
Panel:	Levin, Lin & Chu t*: 0.022 (p-val=0.509)			
	Im, Pesaran and Shin W-stat: -4.023 (p-val=0.000)			

Table 3: Partial adjustment subject to group and individual mean

This table shows adjustment speeds λ conditional on the location to group and country means. Coefficient λ_k of the model $pe_{i,t+k} = \overline{pe}\lambda_k + (1-\lambda_k)\, pe_{i,t} + \varepsilon_{i,t+k}$. λ_c denotes the speed of adjustment for pe ratios when time-series and cross-sectional signals are consistent, λ_o is the speed of adjustment for pe ratios for opposing signals. $\lambda_{>\overline{pe}}$ and $\lambda_{<\overline{pe}}$ denotes speed of adjustment for pe ratios below and above the group mean respectively. $\lambda_{>\overline{pe}}^{>\overline{pe_i}}$ is associated to pe ratios above both group and country specific mean and $\lambda_{<\overline{pe}}^{<\overline{pe_i}}$ below both means. *,** and *** denote significance at the 10%, 5% and 1% level respectively. Standard errors in parenthesis are computed via Seemingly Unrelated Regression (SUR).

Horizon k	$\widehat{\lambda}$	$\widehat{\lambda}_o$	$\widehat{\lambda}_c$	$\widehat{\lambda}_{>\overline{pe}}$	$\widehat{\lambda}_{<\overline{pe}}$	$\widehat{\lambda}_{>\overline{pe}}^{>\overline{pe_i}}$	$\widehat{\lambda}_{<\overline{pe}}^{<\overline{pe_i}}$
Panel A: Developed Countries 1973M01-2014M05							
1 year	0.256***	0.015	0.265***	0.258***	0.254***	0.264***	0.263***
	(0.033)	(0.065)	(0.031)	(0.065)	(0.034)	(0.061)	(0.034)
5 years	0.471***	-0.003	0.487***	0.451***	0.483***	0.490***	0.485***
	(0.054)	(0.282)	(0.041)	(0.110)	(0.037)	(0.092)	(0.036)
Panel B: Emerging/Developing Countries 1991M02-2014M05							
1 year	0.526***	-0.455**	0.543***	0.661***	0.450***	0.663***	0.473***
	(0.053)	(0.194)	(0.049)	(0.068)	(0.060)	(0.066)	(0.061)
5 years	0.843***	0.337*	0.850***	1.105***	0.704***	1.106***	0.711***
	(0.077)	(0.177)	(0.068)	(0.108)	(0.067)	(0.104)	(0.061)
Panel C: Regions 1973M01-2014M05							
1 year	0.171***	-0.043	0.180***	0.174***	0.450***	0.663***	0.473***
	(0.024)	(0.071)	(0.022)	(0.043)	(0.023)	(0.045)	(0.023)
5 years	0.409***	-0.086	0.427***	0.420***	0.704***	1.106***	0.711***
	(0.061)	(0.166)	(0.038)	(0.084)	(0.032)	(0.080)	(0.031)

Table 4: Predictive regressions subject to group and individual mean

This table shows coefficients $\widehat{\beta_k}$ and adj. R^2 for the model conditional on the location to group and country means. Regress $\alpha_k + pe'_{i,t}\ \beta_k + \varepsilon_{i,t\to t+k}$. Samples: *all* denotes the full sample, *opposing* is the sample for pe ratios when time-series and c opposing, *consistent* is the sample for pe ratios with consistent signals. $> \overline{pe}$ and $< \overline{pe}$ denote samples for pe ratios bel mean respectively. $> \overline{pe}, > \overline{pe_i}$ is associated to pe ratios above both group and country specific mean and $< \overline{pe}, < \overline{pe_i}$ belov errors in parenthesis are computed via Seemingly Unrelated Regression (SUR). Asterisks *,**,*** denote significance on th based on the AR(1) model: $r_{i,t\to t+k} = \alpha_k + pe'_{i,t}\ \beta_k + u_{i,t\to t+k}$ with $u_{i,t\to t+k} = \phi_k\ u_{i,t} + \varepsilon_{i,t\to t+k}$.

	all pe		*opposing pe*		*consistent pe*		$> \overline{pe}$		$< \overline{pe}$		$> \overline{pe}, >$
Horizon k	$\widehat{\beta_k}$	R^2	$\widehat{\beta_k}$	R^2	$\widehat{\beta_k}$	R^2	$\widehat{\beta_k}$	R^2	$\widehat{\beta_k}$	R^2	$\widehat{\beta_k}$
Panel A: Developed Countries 1973M01-2014M05											
1 month	-0.006***	0.1%	-0.002	0.0%	-0.006***	0.2%	-0.012**	0.3%	0.002	0.0%	-0.014***
	(0.001)		(0.007)		(0.001)		(0.005)		(0.003)		(0.004)
6 months	-0.056***	1.7%	-0.039	0.2%	-0.057***	2.1%	-0.071***	1.9%	-0.044***	0.4%	-0.077***
	(0.010)		(0.035)		(0.010)		(0.025)		(0.018)		(0.022)
1 year	-0.121***	4.0%	-0.034	0.1%	-0.125***	5.0%	-0.159***	4.1%	-0.104***	1.1%	-0.161***
	(0.021)		(0.073)		(0.020)		(0.049)		(0.030)		(0.013)
5 years	-0.413***	13.3%	0.422	2.3%	-0.442***	18.4%	-0.479***	9.9%	-0.464***	6.9%	-0.458***
	(0.056)		(0.396)		(0.043)		(0.073)		(0.073)		(0.070)
Panel B: Emerging/Developing Countries 1991M02-2014M05											
1 month	-0.015***	0.3%	-0.028	0.1%	-0.014***	0.3%	-0.032**	0.5%	-0.013*	0.1%	-0.031*
	(0.005)		(0.029)		(0.005)		(0.014)		(0.008)		(0.015)
6 months	-0.112***	2.3%	-0.033	0.0%	-0.113***	2.7%	-0.158***	1.7%	-0.131***	1.4%	-0.146***
	(0.028)		(0.121)		(0.026)		(0.066)		(0.050)		(0.072)
1 year	-0.249***	5.6%	0.161	0.2%	-0.254***	6.7%	-0.387***	4.5%	-0.223***	2.1%	-0.377***
	(0.055)		(0.204)		(0.051)		(0.108)		(0.087)		(0.115)
5 years	-0.980***	25.2%	-0.596	1.0%	-0.984***	27.4%	-0.900***	7.4%	-0.539***	5.0%	-0.752***
	(0.164)		(0.367)		(0.152)		(0.274)		(0.150)		(0.273)
Panel C: Regions 1973M01-2014M05											
1 month	-0.004***	0.1%	-0.004	0.0%	-0.004***	0.1%	-0.008***	0.1%	0.000	0.0%	-0.007**
	(0.001)		(0.005)		(0.002)		(0.003)		(0.003)		(0.003)
6 months	-0.039***	1.1%	0.011	0.0%	-0.042***	1.5%	-0.059***	1.3%	-0.046***	0.5%	-0.053***
	(0.009)		(0.025)		(0.009)		(0.014)		(0.017)		(0.016)
1 year	-0.091***	2.9%	0.066	0.2%	-0.098***	4.2%	-0.156***	3.8%	-0.100***	1.4%	-0.149***
	(0.018)		(0.046)		(0.016)		(0.032)		(0.028)		(0.035)
5 years	-0.398***	15.1%	0.575	3.3%	-0.433***	24.8%	-0.632***	14.3%	-0.394***	7.4%	-0.547***
	(0.067)		(0.254)		(0.044)		(0.065)		(0.080)		(0.061)

Table 5: Individual predictive regressions

This table shows predictive regressions for each individual country for 1 month, 3 month and 6 month horizons k. Regression equation: $r_{t\to t+k} = \alpha_k + \beta_k pe_t + \varepsilon_{t\to t+k}$. Coefficients $\widehat{\beta_k}$ and respective standard errors are calculated following Amihud and Hurvich (2004) to correct for the Stambaugh (1999) bias.

	1 month			3 months			6 months		
	$\widehat{\beta_k}$	t	R^2	$\widehat{\beta_k}$	t	R^2	$\widehat{\beta_k}$	t	R^2
Panel A: Developed Countries 1973M01-2014M05									
Australia	-0.009	-1.300	0.003	-0.028	-0.893	0.004	-0.051	-0.773	0.007
Austria	-0.001	-0.067	-0.002	0.005	0.151	-0.001	-0.008	-0.104	-0.001
Belgium	0.006	0.714	-0.002	0.001	0.045	-0.001	-0.039	-0.543	0.007
Canada	-0.004	-0.548	0.001	-0.013	-0.460	0.003	-0.036	-0.652	0.010
Denmark	-0.007	-1.243	0.001	-0.025	-1.582	0.010	-0.076	-2.191	0.043
France	-0.004	-0.517	0.000	-0.029	-0.911	0.007	-0.090	-1.285	0.025
Germany	-0.005	-0.471	0.000	-0.028	-0.695	0.003	-0.066	-0.743	0.008
Hongkong	-0.032	-3.544	0.023	-0.135	-7.270	0.080	-0.248	-8.047	0.117
Ireland	-0.004	-0.450	-0.001	-0.015	-0.463	0.001	-0.043	-0.604	0.008
Japan	-0.004	-0.910	0.000	-0.016	-1.051	0.005	-0.040	-1.004	0.014
Netherland	-0.003	-0.424	0.000	-0.015	-0.699	0.003	-0.041	-1.006	0.012
Singapore	-0.018	-2.004	0.012	-0.072	-2.292	0.033	-0.147	-3.100	0.065
Switzerland	-0.002	-0.381	0.000	-0.012	-0.656	0.001	-0.034	-0.756	0.004
UK	-0.015	-2.082	0.009	-0.053	-2.308	0.028	-0.112	-1.815	0.055
US	-0.010	-2.620	0.006	-0.031	-2.034	0.022	-0.069	-1.939	0.048
Panel B: Emerging/Developing Countries 1991M02-2014M05									
Chile	-0.079	-4.516	0.058	-0.196	-4.947	0.093	-0.346	-5.658	0.139
India	-0.053	-3.135	0.027	-0.149	-3.779	0.068	-0.282	-2.965	0.130
Greece	-0.004	-0.293	0.000	-0.030	-1.138	0.001	-0.065	-1.400	0.009
Indonesia	-0.001	-0.023	0.000	-0.052	-0.869	0.000	-0.174	-1.592	0.019
Korea	-0.013	-0.897	0.003	-0.053	-1.184	0.013	-0.133	-1.466	0.032
Malaysia	0.011	0.837	-0.004	0.000	0.007	-0.002	-0.044	-0.637	0.000
Mexico	-0.018	-0.795	0.000	-0.071	-1.082	0.008	-0.144	-1.261	0.013
Taiwan	-0.015	-1.137	0.011	-0.073	-2.025	0.038	-0.147	-1.785	0.061
Thailand	-0.013	-0.568	0.001	-0.010	-0.169	0.000	-0.008	-0.063	0.000
Sri Lanka	-0.003	-0.442	0.000	-0.022	-1.361	0.006	-0.073	-2.721	0.023
South Africa	-0.050	-3.330	0.019	-0.141	-4.083	0.056	-0.266	-4.752	0.100
Philippines	-0.023	-1.946	0.006	-0.066	-2.119	0.015	-0.135	-3.390	0.028
Panel C: Regions 1973M01-2014M05									
Americas	-0.011	-2.685	0.007	-0.036	-1.982	0.025	-0.079	-2.178	0.055
Asia	-0.003	-0.378	-0.002	-0.005	-0.292	-0.001	-0.022	-0.604	0.003
Australasia	-0.008	-1.211	0.000	-0.025	-0.820	0.003	-0.047	-0.717	0.006
Developed	-0.006	-1.213	0.002	-0.023	-1.427	0.011	-0.057	-1.474	0.026
EMU	0.004	0.413	-0.002	-0.006	-0.177	0.000	-0.039	-0.499	0.004
Europe	-0.005	-0.646	0.001	-0.027	-0.971	0.009	-0.071	-1.127	0.022
Far East	0.000	-0.067	-0.002	-0.007	-0.383	0.000	-0.027	-0.585	0.005
North America	-0.009	-2.335	0.006	-0.030	-1.947	0.021	-0.068	-1.859	0.045
NORCS	0.001	0.195	-0.002	-0.002	-0.063	-0.001	-0.027	-0.498	0.004
Pacific Basin	0.000	0.004	-0.002	-0.006	-0.357	0.000	-0.024	-0.632	0.004
PIIGS	0.001	0.112	-0.002	-0.006	-0.227	-0.001	-0.025	-0.388	0.002
Scandinavia	-0.005	-0.881	0.000	-0.018	-0.916	0.004	-0.056	-1.066	0.019
SE Asia	-0.030	-2.707	0.025	-0.115	-3.710	0.056	-0.219	-4.365	0.091
World	-0.010	-1.985	0.003	-0.027	-1.527	0.013	-0.066	-1.635	0.031

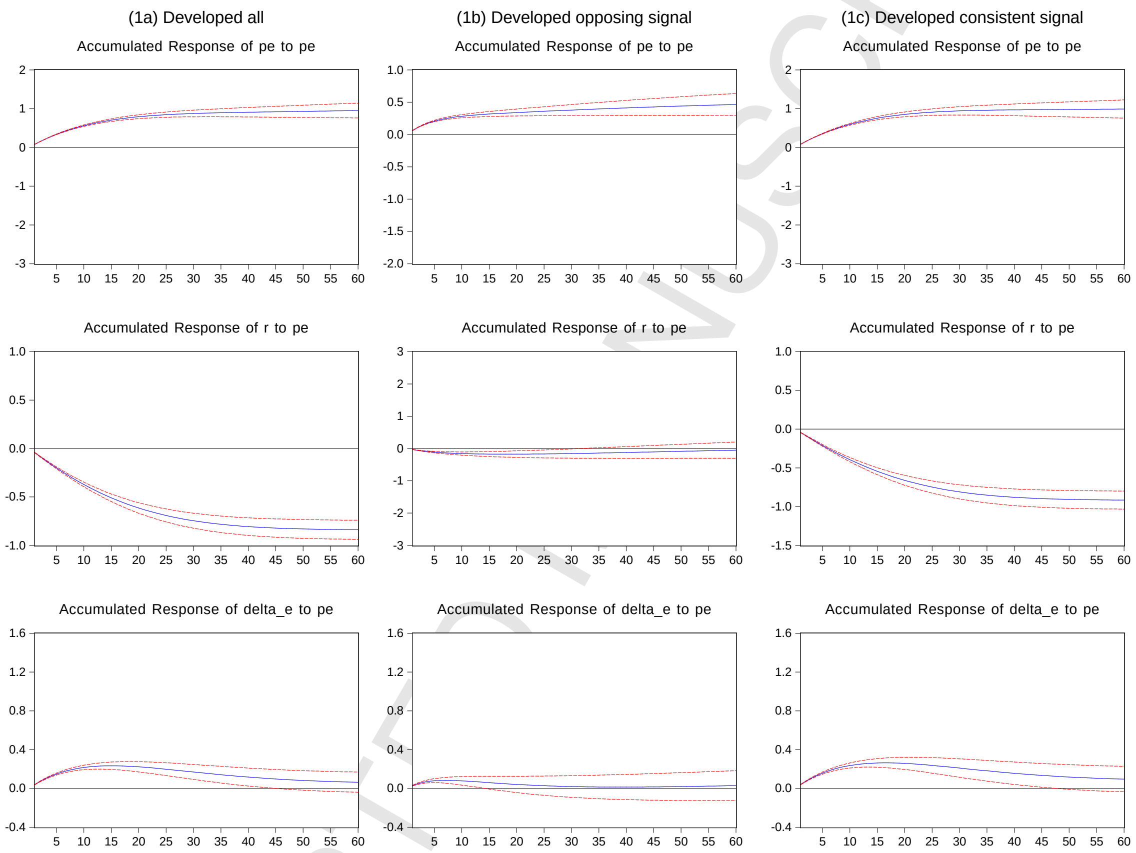

Fig. 3. Accumulated impulse-response functions - developed countries

These graphs show accumulated responses to Cholesky one standard deviation innovations to the *pe* ratio following a first order VAR from Eq. (12). 5 year (60 months) horizon. Vertical axes are in units of the response variable. Standard errors are computed via Monte Carlo simulations, 1000 repetitions.

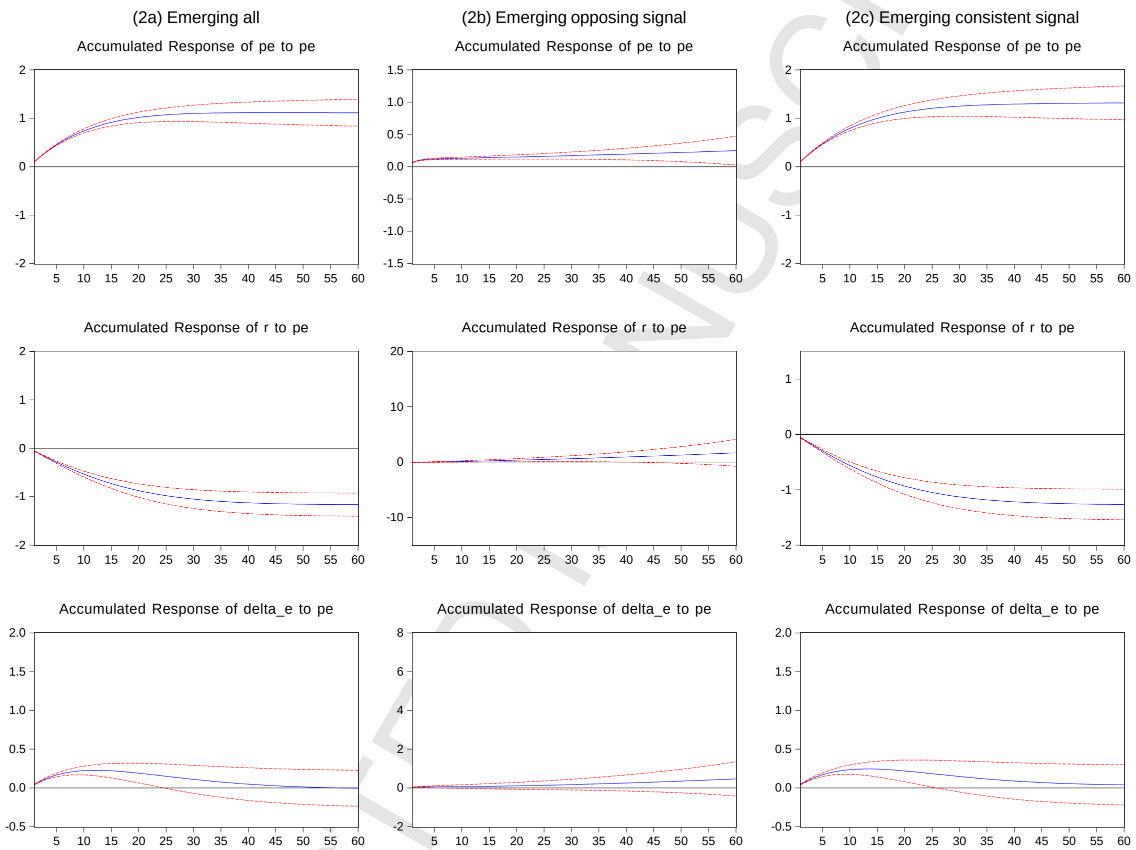

Fig. 4. Accumulated impulse-response functions - emerging countries

These graphs show accumulated responses to Cholesky one standard deviation innovations to the *pe* ratio following a first order VAR from Eq. (12). 5 year (60 months) horizon. Vertical axes are in units of the response variable. Standard errors are computed via Monte Carlo simulations, 1000 repetitions.

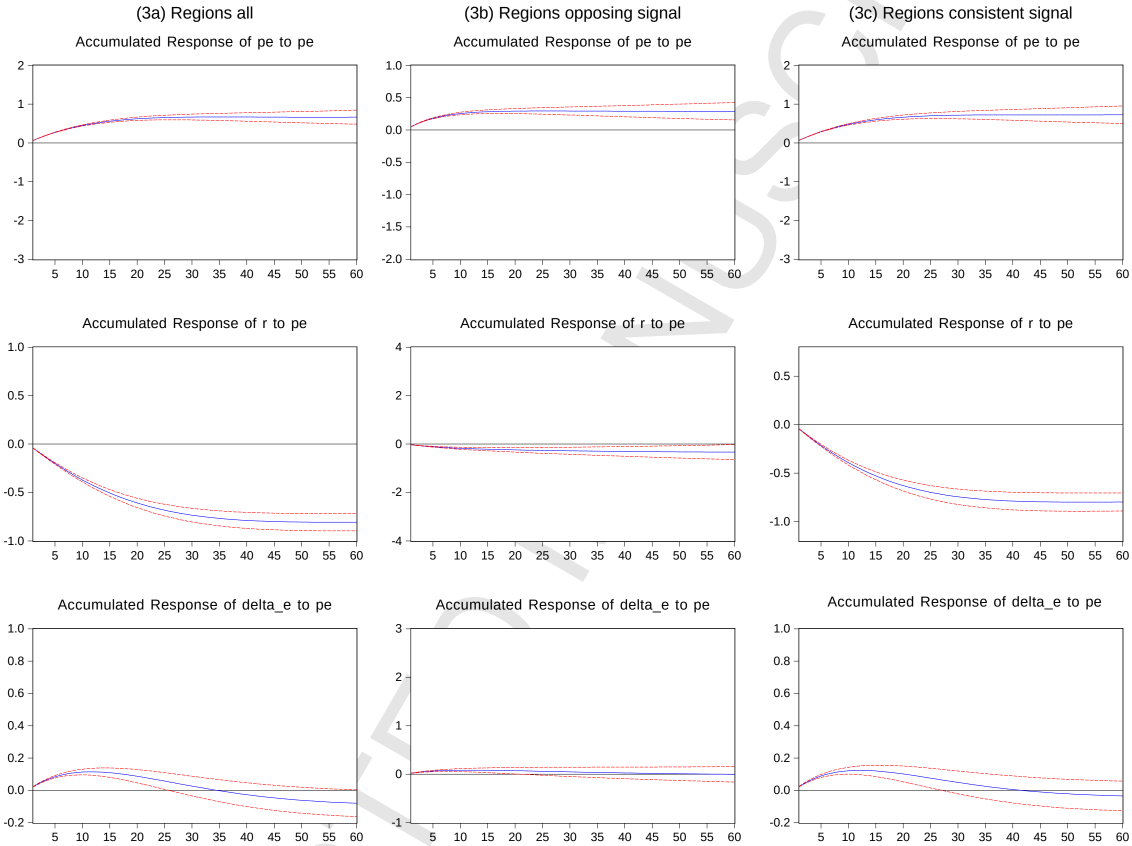

Fig. 5. Accumulated impulse-response functions - regions

These graphs show accumulated responses to Cholesky one standard deviation innovations to the *pe* ratio following a first order VAR from Eq. (12). 5 year (60 months) horizon. Vertical axes are in units of the response variable. Standard errors are computed via Monte Carlo simulations, 1000 repetitions.

Table 6: Out-of-sample results

This table shows out-of-sample statistics for Model 1:

$$r_{i,t\to t+k} = \alpha_k + pe'_{i,t}\,\beta_k + \varepsilon_{i,t\to t+k},$$

and Model 2:

$$r_{i,t\to t+k} = \alpha_k + pe'_{i,t}\,\beta_k + (D^{oos}_{i,t} \times pe_{i,t})'\,\gamma_k + \varepsilon_{i,t\to t+k}.$$

Model 2c denotes the sample for forecasts where the signals were consistent and Model 2o is the sample with forecasts comprising opposing signals only. Out-of-sample test period starts after a 6-year in-sample estimation and $k = 1$ month forecast horizon. Out-of-sample R^2 are provided for a 6 year and 12 year burn-in sample ($R^2_{6\to oos}$ and $R^2_{12\to oos}$ respectively). For the 6 year burn-in sample we provide $DM\ p-val$, $MSFE\ F-stat$ and ΔCER. $DM\ p-val$ denote the p values for the Diebold and Mariano (1995) test. $MSFE\ F-stat$ denotes the F statistic for the test following McCracken (2007). ΔCER denotes the annualized utility gain for a mean-variance investor and risk aversion coefficient of $\gamma = 3$.

Model	MAE [%]	RMSE [%]	TIC [%]	$R^2_{6\to oos}$ [%]	$R^2_{12\to oos}$ [%]	DM p-val	MSFE F-stat	ΔCER [%]
Panel A: Developed Countries								
M 1	4.622	6.428	87.452	0.134	0.426			-1.336
M 2	4.625	6.427	86.842	0.037	0.421	0.031	1.918	-1.044
M 2c	4.077	5.554	80.185	0.761	0.519	-	-	-
M 2o	4.777	6.649	88.710	-0.103	0.060	-	-	-
Panel B: Emerging Countries								
M 1	6.549	9.828	91.771	0.232	0.213			-2.043
M 2	6.534	9.823	91.515	0.319	0.388	0.186	6.122	-0.816
M 2c	5.365	7.721	86.041	0.941	0.722	-	-	-
M 2o	5.858	10.351	92.758	0.223	0.167	-	-	-
Panel C: Regions								
M 1	4.334	5.916	86.541	0.143	0.311			-1.859
M 2	4.336	5.915	85.556	0.172	0.319	0.142	2.014	-1.088
M 2c	3.707	4.989	79.490	1.178	0.884	-	-	-
M 2o	4.475	6.110	88.655	-0.065	-0.101	-	-	-

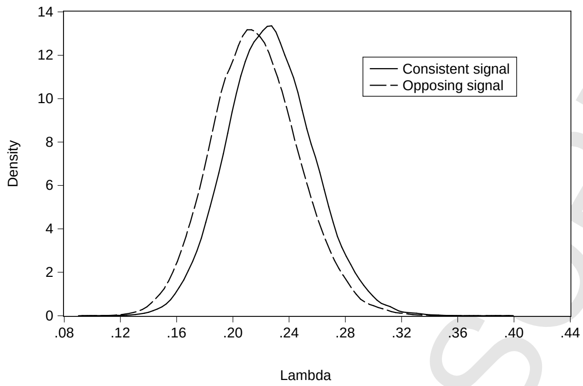

Fig. 6. Distribution of simulated *pe* ratios

This figure shows the distribution of 100,000 repetitions for 15 simulated *pe* ratios. The horizontal axis marks the coefficient λ for the partial adjustment model using the simulated *pe* ratios . Consistent versus opposing signals are calculated for each repetition.

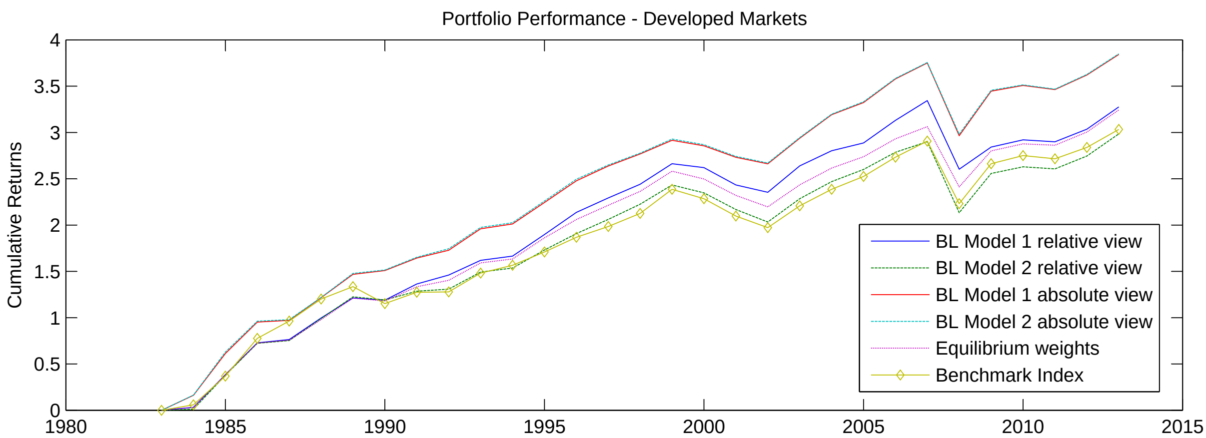

Fig. 7. Portfolio Performance - Developed Markets

This figure shows the cumulative returns for portfolios consisting of developed country indices and a benchmark. Black-Litterman optimized portfolios are estimated with a 10 year in-sample period followed by yearly rebalancing using out-of-sample information.

Table 7: Portfolio Performance Measures

This table shows common portfolio performance statistics for developed and emerging country portfolios. Measures are based on a 10 year pre-estimation and yearly re-balancing according to Black-Litterman tilted weights. M1 denotes Model 1 using the prediction of Eq. (14) as inputs for views $\mathbf{q}$. M2 uses estimates of Eq. (15). Portfolio M2c uses inputs from Model 2 but weights are only tilted according to views when signals of *pe* ratios are consistent. Equilibrium is the market return $\boldsymbol{\pi}$. The benchmark is an index for developed or emerging markets. rel denotes relative views and abs denotes absolute views. Numbers are in %.

	M1 rel	M2 rel	M1 abs	M2 abs	M2c abs	Equilibrium	Benchmark
Panel A: Developed Countries							
$\bar{\mu}$	10.57	9.63	12.39	12.42	12.38	10.46	9.78
Sharpe ratio	32.39	26.3	38.18	38.47	38.33	33.2	27.99
Information ratio	11.29	-2.89	34.42	34.45	34.42	53.9	
Jensens alpha	0.78	-0.16	2.61	2.63	2.59	0.68	
Max drawdown	22.15	26.42	20.95	20.62	20.94	21.23	23.24
Lower partial moment	22.58	22.58	19.35	19.35	19.35	22.58	22.58
Panel B: Emerging Countries							
$\bar{\mu}$	12.73	11.21	17.98	16.08	16.17	11.95	12.29
Sharpe ratio	39.78	31.23	60.23	56.76	57.92	34.61	31.93
Information ratio	5.35	-21.9	21.04	12.37	13.25	-5.74	
Jensens alpha	0.44	-1.08	5.69	3.79	3.82	-0.34	
Max drawdown	38.35	49.91	18.47	23.13	23.70	48.96	47.18
Lower partial moment	30.77	30.77	30.77	23.08	26.92	30.77	38.46

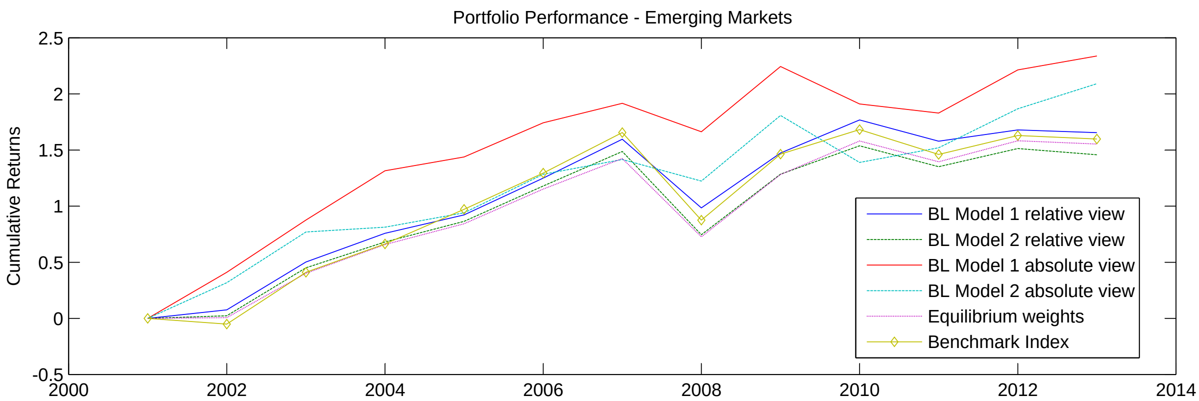

Fig. 8. Portfolio Performance - Emerging Markets

This figure shows the cumulative returns for portfolios consisting of developing country indices and a benchmark. Black-Litterman optimized portfolios are estimated with a 10 year in-sample period followed by yearly rebalancing using out-of-sample information.

Table 8: Partial adjustment subject to group and individual mean: dy and pc

This table shows adjustment speeds λ conditional on the location to group and country means. Coefficient λ_k of the model $dy_{i,t+k} = \overline{dy}\lambda_k + (1 - \lambda_k)\, dy_{i,t} + \varepsilon_{i,t+k}$. λ_c denotes the speed of adjustment for dy ratios when time-series and cross-sectional signals are consistent, λ_o is the speed of adjustment for dy ratios for opposing signals. $\lambda_{>\overline{dy}}$ and $\lambda_{<\overline{dy}}$ denotes speed of adjustment for dy ratios below and above the group mean respectively. $\lambda_{>\overline{dy}}^{>\overline{dy_i}}$ is associated to dy ratios above both group and country specific mean and $\lambda_{<\overline{dy}}^{<\overline{dy_i}}$ below both means. Panels A to C show results for dy, Panel D for pc. *,** and *** denote significance at the 10%, 5% and 1% level respectively. Standard errors in parenthesis are computed via Seemingly Unrelated Regression (SUR).

Horizon k	$\widehat{\lambda}$	$\widehat{\lambda}_o$	$\widehat{\lambda}_c$	$\widehat{\lambda}_{>\overline{dy}}$	$\widehat{\lambda}_{<\overline{dy}}$	$\widehat{\lambda}_{>\overline{dy}}^{>\overline{dy_i}}$	$\widehat{\lambda}_{<\overline{dy}}^{<\overline{dy_i}}$
Panel A: dy Developed Countries 1973M01-2014M05							
1 year	0.137***	-0.084	0.151***	0.209***	0.136	0.113***	0.044
	(0.021)	(0.047)	(0.018)	(0.027)	(0.038)	(0.034)	(0.042)
5 years	0.355***	-0.131	0.389***	0.516***	0.363***	0.431***	0.298**
	(0.058)	(0.135)	(0.044)	(0.062)	(0.085)	(0.079)	(0.119)
Panel B: dy Emerging/Developing Countries 1991M02-2014M05							
1 year	0.185***	-0.087	0.197***	0.170*	0.203	0.083	0.128
	(0.041)	(0.082)	(0.039)	(0.096)	(0.075)	(0.096)	(0.083)
5 years	0.536***	-0.388	0.572***	0.677***	0.451**	0.417***	0.196
	(0.124)	(0.266)	(0.115)	(0.175)	(0.226)	(0.150)	(0.234)
Panel C: dy Regions 1973M01-2014M05							
1 year	0.116***	-0.016	0.125***	0.150***	0.092***	0.170***	0.093***
	(0.020)	(0.059)	(0.017)	(0.024)	(0.025)	(0.024)	(0.024)
5 years	0.357***	-0.299	0.399***	0.424***	0.313***	0.466***	0.356***
	(0.067)	(0.163)	(0.045)	(0.052)	(0.067)	(0.049)	(0.059)
Panel D: pc Developed Countries 1983M01-2014M05							
1 year	0.294***	-0.115	0.319***	0.186***	0.375***	0.208***	0.400***
	(0.039)	(0.064)	(0.035)	(0.041)	(0.048)	(0.043)	(0.043)
5 years	0.571***	-0.286*	0.615***	0.425***	0.679***	0.462***	0.728***
	(0.073)	(0.141)	(0.056)	(0.083)	(0.076)	(0.068)	(0.064)

Table 9: Predictive regressions subject to group and individual mean: dy and pc

This table shows coefficients $\widehat{\beta_k}$ and adj. R^2 for the model conditional on the location to group and country means. Regress
$\alpha_k + dy'_{i,t}\ \beta_k + \varepsilon_{i,t\to t+k}$. Samples: *all* denotes the full sample, *opposing* is the sample for dy ratios when time-series and c
opposing, *consistent* is the sample for dy ratios with consistent signals. $> \overline{dy}$ and $< \overline{dy}$ denote samples for dy ratios bel
mean respectively. $> \overline{dy}, > \overline{dy_i}$ is associated to pe ratios above both group and country specific mean and $< \overline{dy}, < \overline{dy_i}$ belo
to C show results for dy, Panel D for pc. Standard errors in parenthesis are computed via Seemingly Unrelated Regression (
denote significance on the 10%, 5% and 1% level based on the AR(1) model: $r_{i,t\to t+k} = \alpha_k + dy'_{i,t}\ \beta_k + u_{i,t\to t+k}$ with $u_{i,t-}$

	all dy		*opposing dy*		*consistent dy*		$> \overline{dy}$		$< \overline{dy}$		$> \overline{dy}$
Horizon k	$\widehat{\beta_k}$	R^2	$\widehat{\beta_k}$	R^2	$\widehat{\beta_k}$	R^2	$\widehat{\beta_k}$	R^2	$\widehat{\beta_k}$	R^2	$\widehat{\beta}$
Panel A: dy Developed Countries 1973M01-2014M05											
1 month	0.005***	0.1%	0.007	0.1%	0.005***	0.2%	0.010***	0.2%	0.003	0.0%	0.014*
	(0.001)		(0.005)		(0.001)		(0.003)		(0.004)		(0.00
6 months	0.043***	1.4%	0.035	0.2%	0.043***	1.8%	0.091***	1.9%	0.027	0.3%	0.094*
	(0.008)		(0.031)		(0.009)		(0.019)		(0.025)		(0.02
1 year	0.093***	3.4%	0.030	0.1%	0.097***	4.7%	0.148***	2.7%	0.066	0.7%	0.113*
	(0.017)		(0.072)		(0.016)		(0.028)		(0.048)		(0.03
5 years	0.383***	15.9%	-0.438**	4.0%	0.425***	24.9%	0.572***	13.6%	0.394***	6.5%	0.431*
	(0.056)		(0.154)		(0.042)		(0.075)		(0.130)		(0.07
Panel B: dy Emerging/Developing Countries 1991M02-2014M05											
1 month	0.012***	0.3%	-0.007	-0.1%	0.013***	0.5%	0.011	0.0%	0.017**	0.3%	0.000
	(0.003)		(0.012)		(0.003)		(0.008)		(0.008)		(0.00
6 months	0.090***	3.1%	0.002	-0.1%	0.094***	4.1%	0.087*	0.9%	0.097**	1.5%	0.049
	(0.021)		(0.052)		(0.020)		(0.048)		(0.042)		(0.05
1year	0.185***	6.4%	-0.087	0.2%	0.197***	8.9%	0.170*	1.9%	0.203***	3.1%	0.083
	(0.041)		(0.082)		(0.039)		(0.096)		(0.075)		(0.09
5years	0.536***	15.0%	-0.388	1.3%	0.572***	21.0%	0.677***	10.0%	0.451**	4.5%	0.417*
	(0.124)		(0.266)		(0.115)		(0.175)		(0.226)		(0.15
Panel C: dy Regions 1973M01-2014M05											
1 month	0.005***	0.2%	0.002	-0.1%	0.006***	0.2%	0.010***	0.1%	0.008***	0.2%	0.013*
	(0.001)		(0.005)		(0.001)		(0.003)		(0.003)		(0.00
6 months	0.043***	1.6%	0.023	0.1%	0.045***	1.9%	0.087***	1.5%	0.044***	0.9%	0.105*
	(0.009)		(0.019)		(0.008)		(0.024)		(0.015)		(0.03
1year	0.095***	3.6%	0.023	0.0%	0.100***	4.8%	0.145***	2.2%	0.105***	2.1%	0.150*
	(0.018)		(0.029)		(0.016)		(0.038)		(0.027)		(0.04
5years	0.383***	15.6%	-0.262	1.7%	0.424***	24.2%	0.302***	3.3%	0.479***	10.3%	0.233*
	(0.062)		(0.159)		(0.044)		(0.080)		(0.077)		(0.08
Panel D: pc Developed Countries 1983M01-2014M05											
1 month	-0.014***	0.8%	0.005	-0.1%	-0.015***	1.2%	-0.019***	0.4%	-0.017***	0.6%	-0.022*
	(0.003)		(0.007)		(0.002)		(0.006)		(0.005)		(0.00
6 months	-0.086***	4.5%	0.046	0.2%	-0.094***	7.1%	-0.126***	2.6%	-0.107***	3.3%	-0.120*
	(0.017)		(0.036)		(0.014)		(0.037)		(0.029)		(0.03
1 year	-0.164***	7.9%	0.093	0.5%	-0.180***	11.7%	-0.220***	3.5%	-0.179***	5.4%	-0.220*
	(0.032)		(0.056)		(0.026)		(0.072)		(0.046)		(0.07
5 years	-0.410***	13.4%	0.471*	3.5%	-0.460***	21.0%	-0.455**	4.3%	-0.509***	11.0%	-0.455
	(0.091)		(0.228)		(0.073)		(0.191)		(0.128)		(0.19

Table 10: Out-of-sample results for dy and pc

This table shows out-of-sample statistics for Model 1:

$$r_{i,t\to t+k} = \alpha_k + dy'_{i,t}\ \beta_k + \varepsilon_{i,t\to t+k},$$

and Model 2:

$$r_{i,t\to t+k} = \alpha_k + dy'_{i,t}\ \beta_k + (D^{oos}_{i,t} \times dy_{i,t})'\ \gamma_k + \varepsilon_{i,t\to t+k}.$$

Model 2c denotes the sample for forecasts where the signals were consistent and Model 2o is the sample with forecasts comprising opposing signals only. Out-of-sample test period starts after a 6-year in-sample estimation and $k = 1$ month forecast horizon. $DM\ p-val$ denote the p values for the Diebold and Mariano (1995) test. $MSFE\ F-stat$ denotes the F statistic for the test following McCracken (2007). ΔCER denotes the annualized percentage utility gain for a mean-variance investor and risk aversion coefficient of $\gamma = 2, 3, 5$.

Model	MAE [%]	RMSE [%]	TIC [%]	R^2_{oos} [%]	DM p-val	MSE F-stat	$\Delta\text{CER}^{\gamma 2}$	$\Delta\text{CER}^{\gamma 3}$	$\Delta\text{CER}^{\gamma 5}$
Panel A: *dy* Developed Countries									
M1	4.622	6.423	87.254	0.206			1.342	0.921	0.584
M2	4.623	6.429	87.225	0.131	0.121	1.122	-1.071	-0.715	-0.430
M2c	3.948	5.386	82.593	0.611					
M2o	4.697	6.562	88.330	-0.408					
Panel B: *dy* Emerging/Developing Countries									
M1	6.803	10.308	90.628	0.729			-3.741	-0.736	-1.497
M2	6.838	10.338	90.375	0.242	0.192	2.092	-4.027	-0.927	-1.612
M2c	6.266	8.735	89.797	0.333					
M2o	7.052	10.877	93.522	-0.455					
Panel C: *dy* Regions									
M1	4.335	5.920	86.448	0.168			-2.764	-1.841	-1.102
M2	4.341	5.925	86.122	1.397	0.166	1.906	-2.576	-1.716	-1.027
M2c	3.382	4.565	80.343	1.572					
M2o	4.511	6.125	87.937	0.105					
Panel D: *pc* Developed Countries									
M1	4.618	6.430	85.133	0.299			7.505	5.033	3.055
M2	4.620	6.433	84.939	0.242	0.115	1.093	0.534	0.352	0.205
M2c	4.407	6.234	84.180	1.421					
M2o	4.589	6.451	85.394	-0.959					

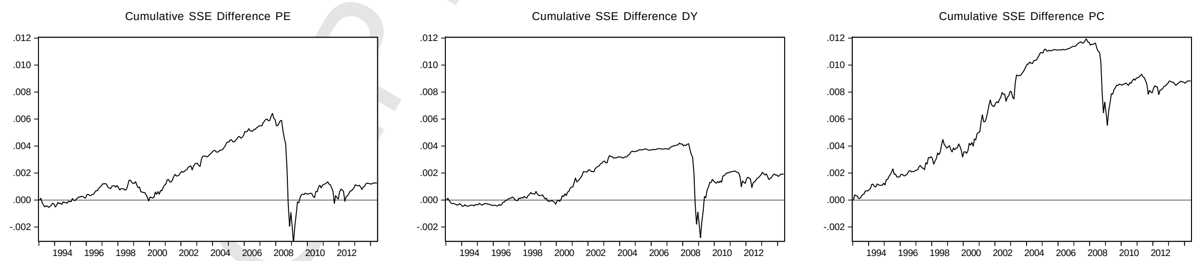

Fig. 9. Cumulative squared prediction errors difference: pe, dy, pc

These figures show the cumulative squared prediction error of the null hypothesis (benchmark mean forecast) minus the cumulative squared prediction error of Model M2c (consistent signals). Calculation as in Goyal and Welch (2008). For comparability of the three predictor variables we choose an out-of-sample period from 1993M01 to 2014M04. The solid lines emphasize means across developed countries.

Table 11: Subsample results

This table contrasts in-sample versus out-of-sample results across sub-periods for *pe*, *dy* and *pc*. In-sample results show regression coefficients ($\beta^c_{k=1m}$) from Eq. (5) for consistent signal ratios and the 1 month horizon. Asterisks *,**,*** denote significance on the 10%, 5% and 1% level (Based on PCSE SUR robust standard errors). R^2 denotes the adjusted R^2. Out-of-sample results are based on portfolios optimized following Black-Litterman with predictions from Model 2 (Eq. (15)). 10 year pre-estimation. Samples consist of developed countries. For the *pc* ratio the out-of-sample period starts in 1994 due to data availability. Numbers for R^2, Sharpe ratio (SR) and Jensen's alpha (α) are in percentage points and the benchmark is the MSCI World.

	in sample		out-of-sample	
	$\beta^c_{k=1m}$	R^2	SR	α
Panel A: *pe*				
83-93	-0.015***	0.89	54.32	4.66
94-04	-0.012***	0.52	22.01	2.47
05-14	0.002	0.07	7.20	0.79
83-14	-0.007***	0.18	38.47	2.59
Panel B: *dy*				
83-93	0.009***	0.53	42.97	2.18
94-04	0.014***	0.86	16.58	1.77
05-14	-0.003	0.09	6.21	0.41
83-14	0.007***	0.26	34.60	1.46
Panel C: *pc*				
-	-	-	-	-
94-04	-0.018***	1.90	45.41	6.70
05-14	-0.007**	0.03	8.04	1.26
94-14	-0.015***	0.82	30.90	3.99

Table 12: Results for simulated data

This table shows coefficients $\widehat{\lambda}$ for the 1 year (k=12) partial adjustment model and $\widehat{\beta}$ for the 1 month predictive regression using simulated pe ratios. Partial adjustment equation: $pe_{i,t+k} = \overline{pe}\lambda_k + (1-\lambda_k)\, pe_{i,t} + \varepsilon_{i,t+k}$. Regression equation: $r_{i,t\to t+1} = \alpha + pe'_{i,t}\,\beta + \varepsilon_{i,t\to t+1}$. Robust standard errors in parenthesis. Asterisks *,**,*** denote significance on the 10%, 5% and 1% level.

	Partial adjustment	Predictive regression	
Signal	$\widehat{\lambda}$	$\widehat{\beta}$	R^2[%]
Consistent	0.225***	-0.002	0.000
	(0.025)	(0.002)	
Opposing	0.219***	-0.003	0.000
	(0.029)	(0.004)	

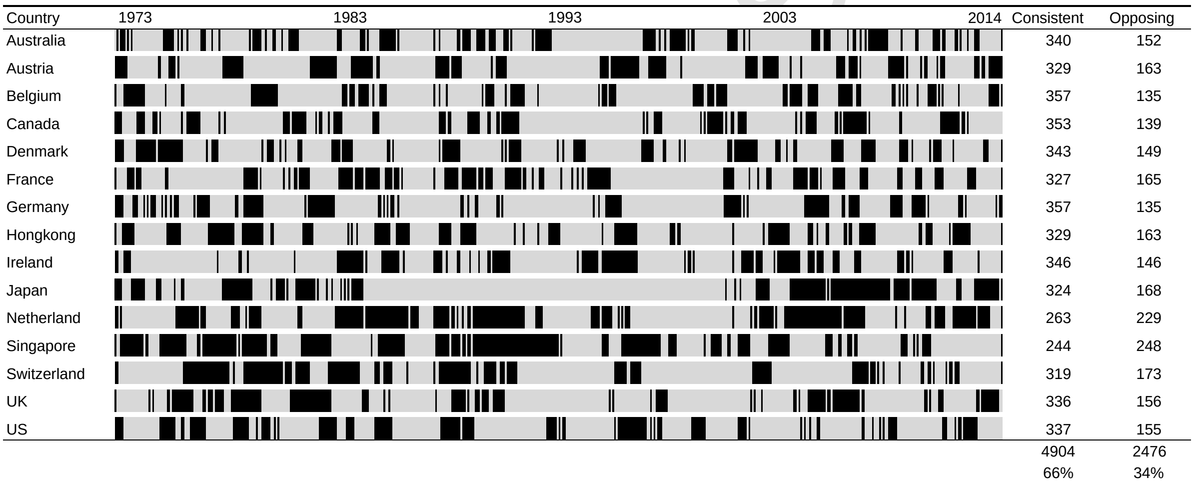

Fig. 10. Consistent versus opposing signals for *pe* ratios

This figure shows summary statistics for developed countries' *pe* ratios. Each black bar highlights a period where signals for current *pe* ratios were opposing (current *pe* ratios are between the global and local mean ratio). Grey bars visualize consistent signals for pe ratios. The two columns on the right show the number of observations for consistent and opposing signals respectively.

1 Highlights

- We find predictability for international stock index returns both in- and out-of-sample
- Joint time-series and cross-section information improves predictive accuracy
- In a Bayesian asset allocation exercise we demonstrates the economic significance of predictability