

**WORKING PAPER** · NO. 2019-92

# Which Investors Matter for Global Equity Valuations and Expected Returns?

*Ralph S.J. Koijen, Robert J. Richmond, and Motohiro Yogo*

JUNE 2019

# Which Investors Matter for Global Equity Valuations and Expected Returns?\*

Ralph S.J. Koijen<sup>†</sup>    Robert J. Richmond<sup>‡</sup>    Motohiro Yogo<sup>§</sup>

June 3, 2019

## Abstract

A large literature in asset pricing decomposes valuation ratios into expected returns and expected growth rates of firm fundamentals to understand why valuation ratios vary across firms and over time. This literature leaves two fundamental questions unanswered: (i) what information do investors attend to in forming their demand beyond prices and (ii) how important are various investors in the price formation process? We use a demand system approach to answer both questions. Empirically, we show that a small set of characteristics explains the majority of variation in a panel of firm-level valuation ratios across countries. We then estimate an international asset demand system using investor-level holdings data in Great Britain and the United States, allowing for flexible substitution patterns within and across countries. We use this framework to measure the contribution of each institutional type in linking characteristics to prices and long-horizon expected returns. Investment advisors, largely driven by their size, are most influential among all institutional types. Conditional on size, hedge funds are the most, and long-term investors (insurance companies and pension funds) are the least influential.

---

\*First draft: March 2019. For comments and discussions, we thank Kent Daniel (discussant), Tarek Hassan (discussant), Carolin Pflueger (discussant), Anna Pavlova, and conference / seminar participants at Boston College, Chicago Booth, Columbia GSB, Notre Dame, NYU Stern, the 2019 Adam Smith Asset Pricing Conference, the 2019 NBER Long-Term Asset Management Conference, and the 2019 UCLA Anderson Fink Center Conference on Financial Markets. We thank Miguel Ferreira and Pedro Matos for discussions regarding the FactSet data.

<sup>†</sup>University of Chicago, Booth School of Business, Center for Economic and Policy Research, and National Bureau of Economic Research (ralph.koijen@chicagobooth.edu).

<sup>‡</sup>New York University (rrichmon@stern.nyu.edu).

<sup>§</sup>Princeton University and National Bureau of Economic Research (myogo@princeton.edu).

A large part of the asset pricing literature centers around two questions. First, how important is variation in expected returns and expected growth rates of fundamentals in explaining fluctuations in asset prices over time and across assets? Second, what explains variation in average returns over time and across assets? The empirical findings of this literature have been influential in guiding modern asset pricing theories, where the focus in recent years shifted towards the role of intermediaries.

The first question is typically answered using predictive regressions of returns or growth rates using valuation ratios, and potentially other variables, as predictors. This methodology is motivated by the present-value identity and follows the seminal work by [Campbell and Shiller \(1988\)](#) in the time series and [Vuolteenaho \(2002\)](#) in the cross-section. In this paper, we are interested in what information investors use so that prices are informative to begin with and how important different investors are for incorporating this information into prices.

To explain differences in average returns, it is common practice to use factor models, such as the capital asset pricing model ([Sharpe 1964](#)), the 3-factor model of [Fama and French \(1992\)](#), and more recently 5-factor models. However, there is disagreement in the literature as to why expected returns vary across assets and over time. To make progress on this question, we show which investors matter in relating characteristics to long-horizon expected returns.

To develop intuition for our empirical strategy, we start with a simple model in which investors use a set of characteristics to forecast a firm's future profitability and to assess its riskiness. Investors form their asset demands as a function of prices, expected future profits, and stocks' riskiness. Investors may disagree on which characteristics are relevant and to what extent they matter for future profits and risk. In equilibrium, prices depend on characteristics, as in hedonic pricing models. The coefficients on the characteristics are a weighted average of the preferences of the investors, where the weights depend on assets under management.

The model guides our empirical analysis. First, we use firm characteristics to explain the cross-sectional variation in valuation ratios. A large literature studies how variation in these valuation ratios relates to expected returns and firm performance (e.g. [Fama and French \(1995\)](#), [Daniel and Titman \(2006\)](#), and [Campbell, Polk, and Vuolteenaho \(2009\)](#)). We show

that a small set of six characteristics related to risk, productivity, and profitability explains the majority of variation in a panel of valuation ratios in the United States (US), the Euro area, Japan, and Great Britain (GB). The residual variation that we cannot explain could be attributable to funding constraints (Brunnermeier and Pedersen 2009), fluctuations in sentiment (Barberis, Shleifer, and Vishny 1998), or potentially omitted characteristics. We also show that the same characteristics predict about a third of the variation in firm’s future profit growth in each of the regions. The coefficients when predicting future profit growth are comparable to the coefficients in the valuation regressions. By the present-value identity, the difference in the coefficients between the valuation and profit growth models, multiplied by the characteristics, is an estimate of long-horizon expected returns.

Next, we study which investors attend to the information embedded in the characteristics so that prices are informative. To this end, we estimate an international asset demand system in which investors’ demands are modeled as a function of prices, characteristics, and latent demand, which captures unobserved demand effects. By imposing market clearing, we can solve for equilibrium asset prices under various counterfactual scenarios. Our modeling approach follows Kojien and Yogo (2019) who estimate a demand system for the cross-section of US equities to study returns instead of valuations.

We extend their model by allowing for cross-country substitution across assets that can be different from the within-country substitution across assets. The cross-country substitution is governed by a single parameter that varies across investor groups. If it equals zero, then cross-country allocations do not depend on valuations or characteristics as one would find in perfectly segmented markets. If it equals one, then the within- and cross-country elasticities are identical. The point estimates range from 0.10 for broker-dealers to 0.32 for investment advisors, implying that the within-country elasticities are higher than the elasticities across countries. However, markets are not perfectly segmented.

We estimate the model using data in GB and in the US from 2006 to 2016. Throughout the paper, we focus on the largest and most liquid stocks by restricting attention to the largest 90% in each market. While we cover the vast majority of each country’s market capitalization, we omit a large number of firms due to the skewness of the firm size distribution (Gabaix 2011). This allows us to better understand the relatively small set of firms which

make up a substantial fraction of economic activity. Also, many of these large firms have global operations, and our estimate of segmentation therefore refers to financial markets as opposed to economic activity more broadly.

A salient fact about the holdings data is that the portfolios of institutional investors (e.g., hedge funds, mutual funds, broker dealers, ...) deviate significantly from market weights. How important are these deviations for the cross section of valuation ratios and expected returns? To answer this question, we start with market clearing and a demand system approach to decompose the cross section of valuation ratios into the sum of deviations of various investors from market weights within their investment universe. In particular, we compute asset prices after changing the demand curves of one particular group of investor types (e.g., hedge funds, mutual funds, broker dealers, ...) to market weights within their investment universe, which is typically a subset of all listed securities. This is equivalent to assigning the asset-weighted average demand curve of all other investors to this group. We then estimate the same valuation regressions in each country using these new prices to see how much the coefficients on the characteristics, and the residuals, change. Combined with a forecasting model for expected future growth rates, this also measures the importance of each investor type in connecting characteristics to long-horizon expected returns.

As we make precise in the motivating structural model, this procedure provides a useful statistical decomposition to measure how much different investors matter for asset prices. This calculation is similar in spirit to [Shiller \(1981\)](#), who calculates what asset prices would look like if discount rates would be constant.<sup>1</sup> In the context of modern asset pricing models, this would be analogous to changing the demand curve of the representative investor so that risk aversion does not change over time in the model of [Campbell and Cochrane \(1999\)](#)<sup>2</sup> or that the demand curve does not shift with time-varying macro-economic risks in the model of [Bansal and Yaron \(2004\)](#). In our model with rich heterogeneity, we only change the demand curve of one type of institutional investors instead of the representative investor.

---

<sup>1</sup>Under the additional, stronger, assumption that the demand system is stable under such large perturbations, this calculation is also informative about what would happen to asset prices if the trend from active to passive continues for a particular institutional type.

<sup>2</sup>A closer analogy to what we do in the paper is the model by [Chan and Kogan \(2002\)](#), which provides a possible micro foundation of the preferences used in [Campbell and Cochrane \(1999\)](#) via heterogeneity in risk aversion across investors. By computing prices assuming all or part of the heterogeneity is removed, one can quantify the importance of risk premium variation for asset prices.

Our emphasis on groups of intermediaries is motivated by a recent literature that highlights the importance of particular groups of intermediaries, such as mutual funds (Basak and Pavlova 2013), broker dealers Adrian, Etula, and Muir (2014); He, Kelly, and Manela (2017), or insurance companies (Ellul, Jotikasthira, and Lundblad 2011), both empirically and theoretically, for asset prices.

Our estimation results point to substantial heterogeneity in demand curves across institutional types. The ultimate impact on asset prices of switching one type to a market-weighted strategy depends (i) the size of the sector and (ii) the demand elasticities with respect to price and characteristics of all other investors.

We find that the impact on valuations when switching different institutions to holding the market portfolio varies substantially across institutional types. We study these changes in valuations both unconditionally and conditional on characteristics. Unconditionally, we find that investment advisors would have the largest impact on valuations if they become passive investors, which is primarily driven by their size. If we normalize the impact by assets under management (AUM), then hedge funds have the largest impact per dollar of AUM. Long-term investors, such as insurance companies and pension funds, have the smallest impact per dollar of AUM. Broker-dealers, which have received significant attention in the recent literature on intermediary asset pricing, hardly matter for prices given their small size.<sup>3</sup> If we zoom in on prices conditional of characteristics, and hence the connection between long-horizon expected returns and characteristics, we find similar results. In particular, hedge funds appear especially important to incorporate information into prices, conditional on their size.

Lastly, in estimating the demand system, we also make a methodological contribution. In general, there are two main challenges when estimating asset demand systems. First, prices are endogenous to demand shocks. To address this endogeneity problem, we use an instrument that is analogous to Koijen and Yogo (2019), and exploits exogenous variation in investment mandates. Second, a large number of investors hold relatively few stocks. One

---

<sup>3</sup>Broker-dealers, and their balance sheet constraints, may matter for asset prices by their ability to extend leverage to other investors such as hedge funds. The traditional stochastic discount factor (SDF) cannot distinguish between the direct impact via their own securities holdings and their indirect impact by providing leverage. We can estimate the importance of the first channel, and find it to be small.

approach is to pool investors by observable characteristics, but this may mask interesting heterogeneity across investors. We instead propose a new shrinkage estimator. We show how to shrink the coefficients on characteristics to the “market’s consensus,” which is similar in spirit to [Black and Litterman \(1991\)](#) in the context of estimating expected returns. The degree of shrinkage depends on the number of stocks held by an investor and vanishes asymptotically. We choose the shrinkage parameters using cross validation to ensure stable demand curve estimates out-of-sample.

## I. Which investors matter for asset prices? A simple model

We present a simple equilibrium model to illustrate how we can measure the importance of various investors in the price formation process. While the model is intentionally stylized, the basic economic insights extend to many asset pricing models, and it helps us to outline the empirical strategy that we follow in subsequent sections.

**Assumptions about beliefs, preferences, and technology** We consider a two-period model with time indexed by  $t = 0, 1$ . There are  $N$  assets indexed by  $n = 1, \dots, N$  and  $I$  investors indexed by  $i = 1, \dots, I$ . The supply of each asset is normalized to one.

Investors have CARA preferences over assets at time 1,

$$\max_{Q_i} \mathbb{E} [-\exp(-\gamma_i A_{1i} + Z_{1i})], \tag{1}$$

where  $Q_i$  is the vector of asset holdings expressed as the number of shares.  $Z_{1i}$  represents other risk factors that impact the investor that can result from benchmarking, outside income, time-varying investment opportunities, et cetera.

The optimization is subject to the intra-period budget constraint

$$A_{0i} = Q_i' P + Q_i^0, \tag{2}$$

where  $P$  denotes the vector of asset prices and  $Q_i^0$  the investment in the outside cash account. The cash account has a price normalized to one and earns a rate of interest that we normalize

to zero. We parameterize the cross-sectional distribution of absolute risk aversion coefficients as  $\gamma_i = \gamma A_{i0}^{-1}$ .<sup>4</sup>

The terminal payoff of the firm is modeled as

$$D_1(n) = B_0(n)\rho_1(n), \quad (3)$$

where  $B_0(n)$  denotes book equity and  $\rho_1(n)$  is the return on equity (ROE) as firms pay out all earnings as dividends in the final period.

This implies that we can write

$$A_{1i} = A_{0i} + Q'_i(D_1 - P) \quad (4)$$

$$= A_{0i} + q'_i(\rho_1 - MB), \quad (5)$$

where  $MB(n) = P(n)/B_0(n)$ , a firm's market-to-book ratio, and  $q_i(n) = Q_i(n)B_0(n)$ . We refer to  $\mathbb{E}_i[\rho] - MB = g_i - MB$  as a measure of (long-horizon) expected return, where  $\mathbb{E}_i[\cdot]$  corresponds to the expectation of investor  $i$ .

Investors may disagree about the expected ROE and its riskiness. We assume a single factor model for ROEs

$$\rho_i = g_i + \beta_i F + \eta, \quad (6)$$

where  $\eta$  and  $F$  are independent, normally distributed with mean zero, and we normalize  $\text{Var}(F) = 1$ .

We assume that investors agree on  $\text{Var}(\eta) = \sigma^2 I$ , but they may disagree on the systematic exposure of stocks to the factor,  $\beta_i$ . In addition, investors may disagree about the expected growth rate,  $g_i$ . We assume that investors agree to disagree and do not revise their beliefs based on asset prices.

In order to estimate the expected growth rate and the riskiness of the firm's future profits, investors rely on public information, "characteristics," and potentially other information, as

---

<sup>4</sup>For this particular choice, the model's implications mimic those of a more standard CRRA utility model, while maintaining the tractability of the CARA-normal model. Our modeling strategy is similar to [Makarov and Schornick \(2010\)](#).

captured by  $\nu^g$  and  $\nu^\beta$ ,

$$g_i(n) = \lambda_i^g x(n) + \nu_i^g(n), \quad (7)$$

$$\beta_i(n) = \lambda_i^\beta x(n) + \nu_i^\beta(n), \quad (8)$$

where  $\nu_i^g$  and  $\nu_i^\beta$  are assumed to be uncorrelated with  $x$ .

We refer to the set  $x$  as the “basis” of characteristics, which includes a constant. The first characteristic is book equity and captures size effects.

To complete the model, we assume for the background risk factor  $Z_{1i} \sim N(\mu_{X_i}, \sigma_{X_i}^2)$  and

$$2\text{Cov}(Z_{1i}, \rho_i(n)) = z_i(n) = \lambda_i^Z x(n) + \nu_i^Z(n). \quad (9)$$

**Investors’ demand curves** The first-order condition for investor  $i$  is given by

$$(g_i - MB) - \gamma_i (\beta_i \beta_i' + \sigma^2 I) q_i + z_i = 0, \quad (10)$$

or equivalently

$$\begin{aligned} q_i &= \frac{1}{\gamma_i} (\beta_i \beta_i' + \sigma^2 I)^{-1} (g_i - MB + z_i) \\ &= \frac{1}{\gamma_i \sigma^2} \left( I - \frac{\beta_i \beta_i'}{\beta_i' \beta_i + \sigma^2} \right) (g_i - MB + z_i) \\ &= \frac{1}{\gamma_i \sigma^2} (g_i - MB) - \frac{c_i}{\gamma_i \sigma^2} \beta_i + \frac{1}{\gamma_i \sigma^2} z_i, \end{aligned} \quad (11)$$

where  $c_i = (\beta_i' \beta_i + \sigma^2)^{-1} \beta_i' (g_i - MB + z_i)$  is a scalar. Hence, an investor’s demand for a given stock depends on its expected return, that is, the expected growth rate of fundamentals relative to the stock’s current valuation, as well as its riskiness and the hedging benefit it provides.

By substituting the assumptions that we made about expected growth rates and the stocks’ riskiness in (7) and (8), we obtain

$$q_i = -\frac{1}{\gamma_i \sigma^2} MB + \frac{1}{\gamma_i \sigma^2} \left( \lambda_i^g - c_i \lambda_i^\beta + \lambda_i^Z \right)' x + \frac{1}{\gamma_i \sigma^2} \left( \nu_i^g - c_i \nu_i^\beta + \nu_i^Z \right). \quad (12)$$

Empirically, we can link portfolio holdings to observable characteristics. However, we cannot tell whether investors attend to a particular characteristics because they view this information as being relevant in forecasting future profits, to assess or hedge risk, or both. Likewise, if an investor deviates from her demand curve, conditional on characteristics, which is the last term in (12), we cannot tell whether this is due to a particular view on expected growth rates or risk.

**Implications for asset prices** By aggregating investors' demands and equating it to supply, we solve for equilibrium asset prices,

$$\sum_i q_i = B, \quad (13)$$

where we use that the supply of each stock is normalized to one and  $q_i(n) = Q_i(n)B(n)$ . Hence, we have

$$\sum_i -\frac{1}{\gamma_i \sigma^2} MB + \frac{1}{\gamma_i \sigma^2} \left( \lambda_i^g - c_i \lambda_i^\beta + \lambda_i^Z \right)' x + \frac{1}{\gamma_i \sigma^2} \left( \nu_i^g - c_i \nu_i^\beta + \nu_i^Z \right) = B \quad (14)$$

that is

$$MB(n) = \left( \sum_i m_i \lambda_i \right)' x(n) + \sum_i m_i \nu_i, \quad (15)$$

where  $e_1$  denotes the first unit vector,<sup>5</sup>  $\lambda_i = \lambda_i^g - c_i \lambda_i^\beta + \lambda_i^Z - \gamma_i \sigma^2 e_1$ ,  $\nu_i = \nu_i^g - c_i \nu_i^\beta + \nu_i^Z$ , and

$$m_i = \frac{\gamma_i^{-1}}{\sum_i \gamma_i^{-1}} = \frac{A_i}{\sum A_i}, \quad (16)$$

given our assumption that  $\gamma_i = \gamma A_i^{-1}$ .

Hence, valuation ratios are connected to characteristics as investors view those characteristics as being relevant to assess risk (via  $\lambda_i^\beta$ ), to forecast future profitability (via  $\lambda_i^g$ ), or for hedging purposes (via  $\lambda_i^Z$ ). Large investors and investors with more extreme views affect

---

<sup>5</sup>Recall that we ordered the characteristics in such a way that book equity is the first characteristic.

prices more and are therefore more important in the price formation process.

**The contribution of investors to price formation** In the context of the model, we explain how we quantify the importance of an investor, or group of investors, in the price formation process. We consider a case in which that investor  $j$  switches to strict market weights, that is,  $Q_j = \theta_j \iota$ , where  $\theta_j$  is chosen to satisfy the budget constraint. Using this demand curve for investor  $j$ , the market clearing condition changes to

$$\sum_{i, i \neq j} q_i + \theta_j B = B,$$

that is,  $\sum_{i, i \neq j} q_i = (1 - \theta_j)B$ . The new market-clearing valuation ratio is

$$MB^{CF}(n) = \lambda_{mb}^{-j} x(n) + \nu_{mb}^{-j},$$

where  $\lambda_{mb}^{-j} = (1 - \theta_j)^{-1} \sum_{i, i \neq j} A_i \lambda_i / \sum_{i, i \neq j} A_i$  and analogously for  $\nu_{mb}^{-j}$ . Hence, the assets are now priced according to the size-weighted average demand curve of all other investors. Indeed, implementing a passive market indexing strategy is equivalent to assigning investor  $j$  the value-weighted demand curve of all other investors.

By comparing  $\lambda_{mb}^{-j}$  to  $\lambda_{mb}$ , we measure the importance of investor  $j$  to incorporate information about fundamentals into prices. Likewise, by comparing  $\nu_{mb}^{-j}$  to  $\nu_{mb}$ , we measure how important investor  $j$  is in explaining the residual in cross-sectional valuation regressions, and hence which investors cause a stock to be a value or growth stock.

Since  $g_R(x) - MB$  is a measure of long-horizon expected returns, where  $g_R$  is the rational expectations forecast of  $\rho_i$  conditional on characteristics  $x$ ,  $g_R(x) \equiv \mathbb{E}[\rho | x]$ , we can also compute the impact of investors on expected returns by comparing  $g(x) - MB$  to  $g_R(x) - MB^{CF}$  and how long-horizon expected returns vary with characteristics.

This calculation is accurate under the assumption that  $g_R(x)$  does not change in the counterfactual, that is, that there are no real effects of some investors switching to passive strategies. To relax this assumption, we would need to augment the model with a production side. While this is an interesting extension to explore in future work, it is beyond the scope of the current paper.

**Intuition behind the econometric strategy** The model can also be used to illustrate the econometric methodology that we propose to estimate the demand system, and in particular what an economically reasonable target is to shrink the estimates of  $\lambda_i$  towards.

We define  $\lambda_{mb} = \sum_i m_i \lambda_i$  and we decompose  $\lambda_i = \lambda_{mb} + \hat{\lambda}_i$ , implying  $\sum_i m_i \hat{\lambda}_i = 0$ . Hence, in the estimation, we shrink  $\hat{\lambda}_i$  to zero or, equivalently, we shrink  $\lambda_i$  to  $\lambda_{mb}$ .

## II. Data and stylized facts

### A. Data

The data on firm fundamentals, stock prices, and portfolio holdings are from FactSet. Details on data construction can be found in Appendix A. While FactSet provides holdings data in many countries, the types of institutions covered varies due to differences in reporting requirements for various institutions. As a result, we only have mutual fund holdings for a large number of countries, which is insufficient for the purposes of this paper.

For the holdings data, we therefore restrict attention to GB and the US. The US holdings data are sourced from 13-F reports and reports to FactSet by individual funds. The GB holdings data are sourced from the UK Share Register (UKSR) in combination with fund holdings. We follow the FactSet methodology, as detailed in the appendix, to combine various sources. We aggregate the holdings to the institutional level by country.

We check the coverage of the data in Figure 2. We plot the total holdings of US firms by UK investors and vice versa for FactSet. As a point of reference we use the IMF's Coordinated Portfolio Investment Survey (CPIS). As the figure illustrates, the FactSet data are highly representative of aggregate cross-border holdings.

We classify holdings into seven types: Investment Advisors (IA), Mutual Funds (MF), Long-term (LT), Hedge Funds (HF), Private Banking (PB), Brokers (BR), and Households (HH). The category Long-term includes primarily insurance companies and pension funds. We use the FactSet classification of investor types to assign institutions to one of the six institutional groups. FactSet classifies an investment firm as an Investment Advisor when the majority of its investments are not in mutual funds and when it is not a subsidiary of a bank, brokerage firm, or insurance company. If the majority of its investments are in mutual

funds, it is instead classified as Mutual Fund.

The household sector is constructed so that total holdings of a company add up to the company’s market capitalization. In rare instances, the total holdings exceed total market cap, in which case we scale the holdings back proportionally. One reason why this may happen is due to short-selling activity, which are not covered in our data as we only have the long positions (Lewellen 2011).

We construct an annual sample of holdings data that begins in 2006 and ends in 2016. While FactSet has data before 2006, the coverage is incomplete. Firm-level fundamentals are from FactSet for the GB, the US, the Euro Area (EU), and Japan (JP). We construct annual firm fundamentals from 2006 until 2016. While we do not have detailed holdings data in the Euro area and Japan, we use data on valuation ratios and fundamentals in these geographies to show that the fundamentals that we use also explain a large share of the variation in valuation ratios and future profits in these geographies.

### *B. Firm granularity*

Table 1 documents firm characteristics along the firm size distribution, as measured by market capitalization. The top two panels are for the US in 2006 and 2016. In 2016, the top 90% of firms of market capitalization is accounted for by only 761 firms. The largest 82 firms already account for 50% of the aggregate market capitalization.

The largest 50% of firms account for only 38% of sales, 33% of investment and employment, yet 63% of net income. This implies that profits are highly concentrated among the largest firms. By comparing the distribution to 2006, we see that these statistics have been fairly stable over time.

In the bottom two panels, we compute similar statistics for the Euro area, Japan, and GB in 2006 and 2016. The patterns are similar, though less extreme, in other countries. In 2016, the top 50% of firms is made up by 44, 85, and 22 firms in the Euro area, Japan, and GB, respectively. At the 90th percentile, these numbers change to 274, 682, and 182. To illustrate the concentration at the top, we list the largest 10 firms in each geography in Table 2.

For the remainder of the paper, we focus on the largest 90% of firms to make sure that

we focus on stocks that are sufficiently liquid, see also [Asness, Moskowitz, and Pedersen \(2013\)](#). We group the bottom 10% into a small-cap portfolio that becomes an outside asset for the investors. Table 1 shows that the economic impact of these firms is small in terms of employment, investment, and profitability.

### *C. Distribution of institutional types*

Figure 1 reports the ownership shares by institutional type. Investment advisors account for the largest share of institutional ownership in both countries, followed by mutual funds, and then pension funds. However, the ownership shares have been fairly stable during the last decade. Table 3 lists the largest investor by type to provide some perspective on the types of institutions that populate the groups.

The distribution of ownership is concentrated as well, see for instance [Azar, Schmalz, and Tecu \(2018\)](#), and this concentration has increased over time ([Itzhak et al. 2018](#)). To illustrate this in our sample, we list the assets under management in 2016 of the largest 10 investors in the US and GB in Table 4. Part of this concentration is driven by the increased popularity of passive indexing strategies, and has resulted in questions regarding the broader impact on market efficiency ([Garleanu and Pedersen 2019](#)). In this paper, we provide a framework to empirically assess how the price formation process is affected if a particular investor or group of investors were to follow a passive indexing strategy instead.

## **III. Cross-sectional valuation regressions: Global evidence**

We show that a small set of characteristics explains the majority of the cross-sectional variation in valuation ratios in the US, GB, the Euro area, and Japan.

### *A. Selection of characteristics*

We consider six characteristics. The first characteristic is log book equity (LNbe), which captures a size effect. To forecast future productivity and profitability, we use the sales-to-book equity ratio, the foreign sales share, the dividend-to-book equity ratio, and the Lerner index. Our use of the foreign sales share is motivated by models such as [Melitz \(2003\)](#) in

which only the more productive firms enter the export market. The Lerner index is a simple measure of markups that is also used in the recent literature on industry concentration and the rise of superstar firms, see for instance [Gutierrez and Philippon \(2017\)](#). The Lerner index is calculated as operating income after depreciation divided by sales. Lastly, we include a stock’s market beta, measured relative to the local market return, as the canonical measure of stock market risk. We cross-sectionally standardize all characteristics by region and year.

### *B. Explaining valuation ratios using characteristics*

We start with the following panel regression of valuation ratios on the characteristics

$$mb_t(n) = a_t + \lambda'_{mb}x_t(n) + \epsilon_t(n), \quad (17)$$

where  $a_t$  are time fixed effects. The results are reported for each of the four regions in [Table 5](#). First, we find that the coefficient on log book equity is negative, while the productivity and markup variables all enter positively. This implies that these characteristics are helpful in predicting future returns or cash flows. Second, we note that the point estimates are quite comparable across regions. The region that deviates somewhat more from the others is Japan. Third, and importantly from our perspective, we can account for a large fraction of the cross-sectional variation in prices. The within year R-squared ranges from 37% (in Japan) to 68% (in GB). In the US, we explain the majority of the variation in the panel of valuation ratios with an R-squared of 52%.

[Table 5](#) shows that the same characteristics explain a substantial fraction of the cross-sectional variation in future profitability, often with similar coefficients in terms of sign and magnitude as in the valuation regressions.<sup>6</sup> We do note that our global sample is quite short, which makes it challenging to accurately estimate expected future earnings. However, our decomposition of the market-to-book-ratio also implies a decomposition of long-horizon expected returns, once combined with a model for expected earnings, via the present-value identities developed in [Cohen, Polk, and Vuolteenaho \(2003\)](#) and [Campbell,](#)

---

<sup>6</sup>[Kacperczyk, Sundaresan, and Wang \(2019\)](#) show that the informativeness of valuation ratios is increasing in the fraction of equity held by foreign investors, in particular in developed economies (see also [Bena et al. \(2017\)](#)).

Polk, and Vuolteenaho (2009). As such, a decomposition of valuations, combined with a model of earnings expectations, yields a decomposition of expected returns. We explore this in more detail in Section VI.

#### IV. A tractable empirical model of the international asset demand system

In this section, we introduce an empirically-tractable asset demand system, which allows for rich heterogeneity in demand curves across investors. The model extends the asset demand system in Kojien and Yogo (2019) by allowing for substitution across countries.

##### A. Notation

There are  $N_c$  assets, indexed by  $n = 1, \dots, N_c$ , in country  $c$ . Lowercase letters denote the logarithm of the corresponding uppercase variables. We denote the characteristics of asset  $n$  in period  $t$  as  $x_t(n)$ . The financial assets are held by  $I$  investors, indexed by  $i = 1, \dots, I$ . One of the investors in each country is a household sector, which holds whatever assets are not held by institutional investors in that country.

##### B. The universe of assets and asset demand

Motivated by the evidence in Section II, we use the top 90% of stocks by market capitalization in each country as the universe of assets. This ensures that our model focuses on pricing the largest firms that capture almost all economic activity among listed firms and that our estimates are not driven by a large number of small and micro-cap firms.

Each investor allocates assets  $A_{it}$  in period  $t$  across assets in its choice sets  $\mathcal{N}_{i,c,t} \subseteq \{1, \dots, N\}$ . An investor's choice set is a subset of assets that the investor considers or is allowed to hold. Restrictions in the choice set may be driven by investment mandates, benchmarking or information frictions that limit an investor's ability to analyze a large universe of stocks (Merton 1987). We refer to stocks within an investor's choice set as inside assets. There is an outside asset in each country. We define the outside asset in a given country as all stocks that are not part of the top 90% and it is indexed by 0 in each country.

An investor's universe is the set of stocks that the investor holds at some point in our

sample from 2006 to 2016, that is  $\mathcal{N}_{i,c} = \cup_{t=1}^T \mathcal{N}_{i,c,t}$ . We denote the number of assets in the investment universe of country  $c$  as  $|\mathcal{N}_{i,c}|$ .

Investors may invest in GB and the US, and can substitute across countries. The portfolio weight of investor  $i$  in stock  $n$  and country  $c$  is

$$w_{i,t}(n, c) = w_{i,t}(n|c)w_{i,t}(c). \quad (18)$$

The first term on the right side,  $w_{i,t}(n|c)$ , is the portfolio weight across stocks in a given country. The second term,  $w_{i,t}(c)$ , is the portfolio weight across countries. We next discuss the models of  $w_{i,t}(n|c)$  and  $w_{i,t}(c)$ , which are guided by the idea that demand elasticities within and across countries may be different.

The portfolio weight on stock  $n$  within country  $c$  is

$$w_{i,t}(n|c) = \frac{\delta_{i,t}(n|c)}{1 + \sum_{m \in \mathcal{N}_{i,c,t}} \delta_{i,t}(m|c)}, \quad (19)$$

where

$$\ln \delta_{i,t}(n|c) = b_{0,i,c,t} + \beta_{0,i,c} m b_t(n) + \beta'_{1,i,c} x_t(n) + \epsilon_{i,c,t}(n), \quad (20)$$

and  $b_{0,i,c,t}$  are investor-country-time fixed effects. An investor's demand depends on the log market-to-book ratio, firm characteristics, and latent demand,  $\epsilon_{i,c,t}(n)$ . Latent demand captures investor  $i$ 's demand beyond characteristics. This demand curve is motivated by the simple model in Section I, but is empirically more realistic as portfolio holdings are log-normally distributed.

We normalize the mean of latent demand  $\epsilon_{i,c,t}(n)$  to zero so that the intercept  $b_{0,i,c,t}$  in equation (20) is identified. This implies that the error terms average to zero for a given investor across stocks in each year, but the error terms do not necessarily average to zero across investors for a given stock. Indeed, the residual variation in market-to-book ratios beyond characteristics is due to latent demand, see equation (15).

To specify the allocation across countries,  $w_{i,t}(c)$ , we define  $\zeta_{i,t}(c) = 1 + \sum_{m \in \mathcal{N}_{i,t}} \delta_{i,t}(m|c)$ , which is the inverse of the fraction invested in the outside asset,  $\zeta_{i,t}(c) = \frac{1}{w_{i,t}(0|c)}$ . Intuitively,

when  $\zeta_{i,t}(c)$  is large, inside assets are relatively attractive to investor  $i$ , compared to the outside asset, in country  $c$ . This happens when an investor considers the prices to be low relative to fundamentals and latent demand. In this case, the investor may also want to reallocate assets from the foreign country to country  $c$ .

Following this intuition, we model the portfolio weight of country  $c$  as

$$w_{i,t}(c) = \frac{\zeta_{i,t}(c)^{\psi_{1,i}} \delta_{i,t}(c)}{\zeta_{i,t}(US)^{\psi_{1,i}} \delta_{i,t}(US) + \zeta_{i,t}(UK)^{\psi_{1,i}}}, \quad (21)$$

where

$$\ln \delta_{i,t}(US) = \psi_{0,i} + \epsilon_{i,t}^{\psi}, \quad (22)$$

and  $\delta_{i,t}(UK) = 1$ , which is a normalization. The model in (21) implies that the country weight is increasing in  $\zeta_{i,t}(c)$ , that is, the relative attractiveness of inside assets in country  $c$ . Our model of portfolio weights is a nested logit model.

It is useful to consider two special cases to gain intuition. When  $\psi_{1,i} = 1$ , the model collapses to a standard logit model

$$w_{i,t}(n, c) = \frac{\delta_{i,t}(n|c) \delta_{i,t}(c)}{\sum_{k=1}^2 \sum_{m=0}^N \delta_{i,t}(m|k) \delta_{i,t}(k)}, \quad (23)$$

and the elasticity of substitution within and across countries is identical. This implies that the equity markets in GB and the US are perfectly integrated.

When  $\psi_{1,i} = 0$ ,

$$w_{i,t}(n, c) = \frac{\delta_{i,t}(n|c)}{\sum_{m=0}^N \delta_{i,t}(m|c)} \frac{\delta_{i,t}(c)}{\sum_{k=1}^2 \delta_{i,t}(k)}, \quad (24)$$

the allocation across asset classes depends only on  $\psi_{0,i} + \epsilon_{i,t}^{\psi}$ . In this case, the equity markets in both countries are segmented and the relative allocation does not respond to prices, characteristics or latent demand in either country. Empirically, we expect the willingness of investors to substitute within a country to be higher than across countries, and we therefore

expect  $\psi_{1,i} \in (0, 1)$ .<sup>7</sup>

### C. Market clearing

We complete the model with the market clearing condition for each asset  $n$  in country  $c$

$$ME_t(n, c) = \sum_{i=1}^I A_{i,t} w_{i,t}(n, c), \quad (25)$$

that is, the market value of shares outstanding must equal the asset-weighted sum of portfolio weights across all investors. In solving for equilibrium asset prices, we follow the literature on asset pricing in endowment economies (Lucas 1978) and assume that shares outstanding and the characteristics are exogenous.

Koijen and Yogo (2019) show that the following assumption is a sufficient condition for both individual and aggregate demand to be downward sloping.

ASSUMPTION 1. The coefficient on log market equity satisfies  $\beta_{0,i,c} < 1$  for all investors and in each country.

We impose this condition in estimating the demand system, which we discuss in the next section. Readers may choose to skip this section the first time reading the paper and directly go to Section VI that presents the empirical results.

## V. Estimating the asset demand system

We discuss in this section how we estimate international asset demand system as summarized by equations (19)-(20) and (21)-(22). We first discuss an estimator of the demand system within each country that accounts for the fact that investors may hold relatively concentrated portfolios. We then discuss the instrument that we use for prices, and we conclude by developing a simple estimator for the cross-country model.

---

<sup>7</sup>Theoretically, there is no upperbound on  $\psi_{1,i}$  and values above one would imply that investors are more willing to substitute across countries compared to within countries.

A. *A ridge instrumental variables estimator of the within-country demand curves*

We write the demand equation of investor  $i$  in a given country as

$$\ln \left( \frac{w_{i,t}(n)}{w_{i,t}(0)} \right) = b_{0,i,t} + \beta_{0,i} mb_t(n) + \beta'_{1,i} x_t(n) + \epsilon_{i,t}(n), \quad (26)$$

where we omit country subscripts to simplify the notation as we focus on the within-country demand system.

To estimate the demand system within a country, there are two main challenges. First, many investors hold relatively concentrated portfolios consisting of a fairly small number of stocks. This problem is amplified by the fact that (i) we consider the top 90% of stocks in GB and the US and (ii) there are fewer stocks in GB. Earlier literature, e.g. [Kojien and Yogo \(2019\)](#), pool holdings across investors, but this may mask interesting heterogeneity. In this paper, we propose a shrinkage estimator that still provides reasonable estimates, even for small portfolios. We discuss this estimator in the remainder of this section.

The second challenge in estimating the demand system is that prices are endogenous to latent demand, that is,  $\mathbb{E}[\epsilon_{i,t}(n)mb_t(n)] \neq 0$ . The exogeneity condition may be violated for large investors that directly impact prices, but also when latent demand is correlated across investors that are all small (or even atomistic). For instance, this can happen when a firm introduces a new product that is not yet reflected in fundamentals but investors update their forecasts of corporate profits in response. Empirically, latent demand follows a factor model and the common components are correlated with prices. To address this endogeneity problem, we introduce an instrument in the next section and we assume for now that an instrument  $z_{i,t}(n)$  exists that is relevant,  $\mathbb{E}_t[z_{i,t}(n)mb_t(n)] \neq 0$ , and exogenous,  $\mathbb{E}_t[z_{i,t}(n)\epsilon_{i,t}(n)] = 0$ .

The estimator that we propose is a standard two-stage least squares estimator (2SLS), with the exception that we shrink the second-stage estimates of the coefficients on characteristics,  $\beta_{1,i}$ , towards an economically-motivated target. We implement the shrinkage estimator by adding a ridge penalty ([Hoerl and Kennard 1970](#)) to the second-stage estimator, which vanishes asymptotically so that the penalty has no impact for investors who hold enough stocks in their portfolios to reliably estimate their demand curves. We use a cross-validation

procedure to determine the optimal penalty function and how quickly it should vanish as the number of holdings increases. We first determine the shrinkage target and then discuss how to adjust the 2SLS procedure to add the shrinkage penalty.

We propose to shrink the coefficients to the “market’s valuation of characteristics.” The idea to use the market’s assessment of moments of asset prices as a prior or shrinkage target has been explored successfully before by [Black and Litterman \(1991\)](#) to estimate expected returns.<sup>8</sup> Instead of recovering expected returns, we are interested in the market’s valuation of characteristics.

To derive the shrinkage target, consider a setting with a single representative investor that has the same demand curve as in (19)-(20),

$$w_t(n) = \frac{\delta_t(n)}{1 + \sum_m \delta_t(m)},$$

where  $\ln \delta_t(n) = b_t + \beta_0 mb_t(n) + \beta_1' x_t^*(n) + \epsilon_t(n)$ . We standardize all the characteristics cross-sectionally, denoted by  $x_t^*(n)$ , and order them such that the first characteristic is log book equity,  $be_t(n)$ .

Market clearing implies

$$ME_t(n) = A_t w_t(n),$$

and hence

$$mb_t(n) = aw_t - be_t(n) + \beta_0 mb_t(n) + \beta_1' x_t^*(n) + \epsilon_t(n),$$

where  $aw_t = \ln(A_t w_t(0))$ . This implies

$$mb_t(n) = c_{mb,t} + \gamma' x_t^*(n) + u_t(n), \tag{27}$$

where  $e_1$  denotes the first unit vector,  $\sigma_b$  the cross-sectional standard deviation of log book equity,  $\mu_b$  its cross-sectional mean,  $c_{mb,t} = (aw_t - \mu_b)/(1 - \beta_0)$ ,  $\gamma = (\beta_1 - e_1 \sigma_b)/(1 - \beta_0)$ , and  $u_t(n) = \epsilon_t(n)/(1 - \beta_0)$ . We can estimate all parameters in (27) using cross-sectional

---

<sup>8</sup>[Black and Litterman \(1991\)](#) propose a Bayesian estimator of expected returns. The prior is based on the “market’s view of expected returns,” which they uncover by inverting the mean-variance equation (assuming we know the covariance matrix) and setting the weights equal to the market weights. They then combine this prior with experts’ views of expected returns to obtain the posterior estimate of expected returns.

valuation regressions, where  $c_{mb,t}$  is a time fixed effect, just as in (17).

We can plug (27) back into (26)<sup>9</sup>

$$\begin{aligned}
\ln\left(\frac{w_t(n)}{w_t(0)}\right) &= b_{0,t} + \beta_0 mb_t(n) + \beta'_1 x_t(n) + \epsilon_t(n) \\
&= b_{0,t} + \beta_0 (c_{mb,t} + \gamma' x_t^*(n) + u_t(n)) + (1 - \beta_0)\gamma' x_t^*(n) + (be_t(n) - \mu_b) + \epsilon_t(n) \\
&= \widehat{b}_{0,t} + \beta_0 u_t(n) + \gamma' x_t^*(n) + be_t(n) + \epsilon_t(n),
\end{aligned} \tag{28}$$

where  $\widehat{b}_{0,t} = b_{0,t} + \beta_0 c_{mb,t} - \mu_b$ .

This derivation suggests a natural shrinkage target of  $\gamma$  for the characteristics when regressing  $\ln\left(\frac{w_t(n)}{w_t(0)}\right) - be_t(n)$  on characteristics and  $u_t(n)$ . Unfortunately, we cannot learn from the cross-sectional regressions about  $\beta_0$ .<sup>10</sup> In the presence of heterogenous investors as in Section I, which is the empirically relevant case, we obtain an equation as in (28) but with  $\epsilon_t(n)$  replaced with  $\epsilon_{it}(n)$ . In this case,  $u_t(n)$  is the size-weighted average of demand shocks.

We use the following estimation procedure to estimate the within-country asset demand system:

1. Estimate the valuation model in (27) and compute the valuation residual  $u_t(n)$ .
2. Form “normalized demand” as

$$y_{it}(n) = \ln\left(\frac{w_{it}(n)}{w_{it}(0)}\right) - \gamma' x_t^*(n) - be_t(n). \tag{29}$$

3. Estimate the demand curve of investor  $i$

$$y_{it}(n) = c_{0it} + \beta_{0i} u_t(n) + c'_{1i} x_t^*(n) + \epsilon_{it}(n), \tag{30}$$

using the instruments while shrinking  $c_{1i}$ , the deviations from the market’s consensus, toward zero. We leave  $\beta_0$  unconstrained unless we estimate it to be larger than one. In that case, we shrink it to one.

We now elaborate on how we shrink the estimates in the third step. We start from the

---

<sup>9</sup>We drop subscripts  $i$  as we are considering a representative investor in developing the shrinkage target.

<sup>10</sup>Indeed, substituting  $u_t(n) = \frac{\epsilon_{it}(n)}{1-\beta_0}$  implies that  $\beta_0 u_t(n) + \epsilon_{it}(n) = \epsilon_{it}(n)$ .

standard instrumental-variables estimator and implement it as a 2SLS estimator. In the first stage, we regress the endogenous variable,  $u_t(n)$ , on the instrument and the characteristics,<sup>11</sup>

$$u_t(n) = \alpha_{0,i,t} + \alpha_{1,i}z_{i,t}(n) + \alpha'_{2,i}x_t^*(n) + \nu_{i,t}(n),$$

with the fitted values denoted by  $\hat{u}_{i,t}(n) = \hat{\alpha}_{0,i} + \hat{\alpha}_{1,i}z_{i,t}(n) + \hat{\alpha}'_{2,i}x_t^*(n)$ .

In a standard 2SLS procedure, we estimate the coefficients using the least-squares objective

$$(\hat{\beta}_{0,i}, \tilde{c}'_{1,i}, \tilde{c}'_{0,i,t})' = \arg \min_{\beta_{0,i}, c_{1,i}, c_{0,i,t}} \frac{1}{N_i} \sum_{n,t} (y_{i,t}(n) - c_{0,i,t} - \beta_{0,i}\hat{u}_{i,t}(n) - c'_{1,i}x_t^*(n))^2,$$

and recall that  $N_i$  denotes the number of positions held by investor  $i$  across all years.

We modify this objective by adding two penalties. The first penalty that we add,  $\lambda_{0,i}(\beta_{0,i} - 1)^2$ , ensures that  $\beta_{0,i} < 1$ . If the estimate of  $\beta_{0,i}$  already satisfies this constraint with  $\lambda_{0,i} = 0$ , then we set  $\lambda_{0,i}$ . Otherwise, we set it to a large value to ensure that  $\beta_{0,i} = 1$ . As it varies across investors whether the unconstrained estimate of  $\beta_{0,i}$  is below one, the penalty  $\lambda_{0,i} \in \{0, \infty\}$  is indexed by  $i$ .

The second penalty,  $\lambda_1 N_i^{-\xi} c'_{1,i} c_{1,i}$ , shrinks the estimates of  $c_{1,i}$  to a vector of zeros. The importance of the penalty is controlled by  $\lambda_1 N_i^{-\xi}$ , with  $\xi > 0$ , which implies that the penalty becomes less important as investors hold more stocks.  $\xi$  controls the speed at which the penalty vanishes asymptotically.

If we add the two penalties to the standard second-stage objective, we obtain

$$\begin{aligned} (\hat{\beta}_{0,i}, \tilde{c}'_{1,i}, \tilde{c}'_{0,i,t})' = & \hspace{15em} (31) \\ \arg \min_{\beta_{0,i}, c_{1,i}, c_{0,i,t}} \frac{1}{N_i} \sum_{n,t} & (y_{i,t}(n) - c_{0,i,t} - \beta_{0,i}\hat{u}_{i,t}(n) - c'_{1,i}x_t^*(n))^2 + \lambda_{0,i}(\beta_{0,i} - 1)^2 + \lambda_1 N_i^{-\xi} c'_{1,i} c_{1,i}, \end{aligned}$$

where  $\lambda_{0,i}, \lambda_1, \xi \geq 0$ . When  $\lambda_{0,i} = \lambda_1 = 0$ , we obtain the standard instrumental-variables estimator.

---

<sup>11</sup>Although  $u_t(n)$  is, by construction, uncorrelated with characteristics, the instrument may not be and  $\alpha_{2i}$  is non-zero as a result.

We can analytically solve the optimization problem in (31) and solution is given by

$$(\widehat{\beta}_{0,i}, \widehat{\mathcal{C}}'_{1,i}, \widehat{\mathcal{C}}'_{0,i,t})' = \left( N_i^{-1} \sum_{n,t} [\widehat{x}_t(n) \widehat{x}_t(n)'] + \Lambda_i \right)^{-1} \left( N_i^{-1} \sum_{n,t} [\widehat{x}_t(n) y_{i,t}(n)] + \Lambda_i \beta_c \right),$$

where  $\Lambda_i = \text{diag}(\lambda_{0,i}, \lambda_1 N_i^\xi \iota, 0_T)$ ,  $\beta_c = (1, 0_{1 \times K}, 0_T)'$ ,  $\widehat{x}_t(n) = (\widehat{u}_t(n), x_t^*(n), d_t)$ .  $d_t$  is a set of time-fixed effects that we do not shrink. With the closed-form solution in hand, estimating the demand system is fast and can be done in a matter of minutes for all investors.

To complete the estimation procedure, we need to determine  $(\lambda_1, \xi)$ . As is common practice in the machine-learning literature, we choose both parameters using cross-validation. In particular, we split the holdings randomly in half for each investor by year. We then estimate the model on one sample for each investor and compute the mean-squared error on the left out sample. Figure 3 illustrates the mean-squared error for various combinations of  $(\lambda_1, \xi)$ . The mean-squared error is minimized for  $(\lambda_1, \xi) = (2.5, 0.7)$ .

### B. Construction of the instrument

As discussed before, we cannot simply estimate investors' demand curves using ordinary least squares, as latent demand is likely correlated with prices. To construct an instrument, we follow [Kojien and Yogo \(2019\)](#) and use exogenous variation in investors' investment mandates to generate exogenous variation in demand.

The key economic assumption is that the set of stocks that an investor holds over time is exogenous. Investors may drop certain stocks in a particular year as a result of variation in latent demand, that is, that is,  $\epsilon_{i,t}(n)$  and  $\mathcal{N}_{i,t}$  are correlated, but the boundaries of the investment universe,  $\mathcal{N}_i = \cup_{t=1}^T \mathcal{N}_{i,t}$ , are assumed to be exogenous. The boundaries of an investor's investment universe in case of institutional investors is typically determined by investment mandates.

Within the universe of mandates, the actual portfolio holdings are endogenous. As instruments, we compute the counterfactual prices as if investors hold a  $1/(1 + |\mathcal{N}_j|)$  portfolio,

excluding the investor’s own assets and the assets of the household sector

$$z_{i,t}(n) = \log \left( \sum_{j \neq i, HH} A_{j,t} \frac{1_j(n)}{1 + |\mathcal{N}_j|} \right).$$

### C. Estimating the cross-country demand curves

To complete the model estimation, we estimate  $\psi_{0,i}$  and  $\psi_{1,i}$  in (21)-(22) that determine the cross-country demand curves. The model implies

$$\begin{aligned} \ln \left( \frac{w_{i,t}(US)}{w_{i,t}(GB)} \right) &= \psi_{0,i} + \psi_{1,i} \ln \left( \frac{\zeta_{i,t}(US)}{\zeta_{i,t}(GB)} \right) + \epsilon_{i,t}^\psi \\ &= \psi_{0,i} - \psi_{1,i} \ln \left( \frac{w_{i,t}(0|US)}{w_{i,t}(0|GB)} \right) + \epsilon_{i,t}^\psi \end{aligned} \quad (32)$$

where we use  $\zeta_{i,t}(c) = \frac{1}{w_{i,t}(0|c)}$ . Equation (32) highlights that  $\psi_{0,i}$  controls the average allocation to GB relative to the US.  $-\psi_{1,i}$  is the elasticity of the total GB-US share with respect to the GB-US share in the respective outside assets. Intuitively,  $\psi_{1,i}$  measures how much an investor would shift assets from GB to the US when, within GB, the investor shifts from inside assets to the outside asset.

We estimate (32) using a pooled regression by investor type with investor fixed effects. The investor fixed effects provide an estimate of  $\psi_{i,0}$  and that we use within-investor variation in the GB-US share over time to identify  $\psi_1$ .

## VI. Empirical results

We report the estimation results in Section A. In Section B, we define the counterfactual to determine how much different investors matter for asset prices. In Section C, we compute the counterfactual using the earlier estimates. Lastly, we show how different investors impact the link between characteristics and either valuations or expected returns in Section D.

### A. Estimation results

Table 6 summarizes the estimation results for the US (top panel) and GB (middle panel). In the bottom panel, we report the estimates of the cross-country allocation model, that is,  $\psi_{1,i}$ .

The columns correspond to different institutional types and we report the AUM-weighted average of the coefficients in the table.

In the top and middle panel, the coefficient on the log market-to-book ratio,  $\beta_{0,i,c}$ , captures the demand elasticity with respect to price. Lower values correspond to demand curves that are more sensitive to prices. We find that hedge funds are the most elastic institutional investors, while long-term investors (pension funds and insurance companies) are among the least elastic investors in both countries.

The remaining coefficients reflect the response of demand to characteristics. In interpreting the estimates, it is important to keep in mind that we report estimates of  $c_{1,i}$ , see (30), which are deviations from the market's consensus. The overall response of an investor's demand to characteristics is given by  $\gamma + c_{1,i}$ , where  $\gamma$  is the sensitivity of the log market-to-book ratio to characteristics, see (27). For instance, the positive coefficient on foreign sales in the case of mutual funds and the negative coefficient in the case of hedge funds implies that mutual funds tend to overweight firms that export more while hedge funds underweight these firms, all else equal.

To interpret the magnitude of the coefficients, note that for any two investors with  $w_{i,t}(n|c) = w_{j,t}(n|c) = w_t(n|c)$ , the differential response to a change in characteristics is given by

$$\frac{\partial w_{i,t}(n|c)}{\partial x_t^*(n)} - \frac{\partial w_{j,t}(n|c)}{\partial x_t^*(n)} = w_t(n|c)(1 - w_t(n|c))(c_{1,i} - c_{1,j})' x_t^*(n).$$

Hence, suppose two investors invest 10% in a stock and the difference in coefficients is 0.1, which is a typical order of magnitude given Table 6. Then if a particular characteristic increases by one standard deviation, the differential response in portfolio weights is given by  $0.1 \times 0.9 \times 0.1 \times 1.0 = 0.9\%$ . Indeed, for a reasonably diversified investor, so that  $w_t(n|c)(1 - w_t(n|c)) \simeq w_{i,t}(n|c)$ , the difference in coefficients  $c_{1,i} - c_{1,j}$  measures the difference in semi-elasticities with respect to one standard deviation change in characteristics.

The estimates imply that investors disagree in particular about the valuation of dividend-to-book equity, log book equity, and foreign sales in both countries.

In the bottom panel of Table 6, we report the estimates of  $\psi_{1,i}$  in equation (21). Recall that  $\psi_1$  equal to one corresponds to the same substitution patterns across and within coun-

tries, while  $\psi_1$  equal to zero implies that the cross-country shares are insensitive to relative prices and characteristics. The estimates range from 0.10 for broker-dealers to 0.32 for investment advisors. While this does imply that substitution across countries is lower than within countries, markets are not fully segmented.

### *B. Measuring the importance of investors for asset prices*

All investors are marginal in pricing assets, but changes in demand of some investors are more important than changes in demand of other investors. To measure the relative contribution, we compute prices under the assumption that one group of institutional investors switches to holding the market portfolio. As explained in Section I, this implicitly implies that we assign the size-weighted demand curve of all other investors to this group.

The impact on asset prices of switching one particular group of institutional investors to market weights is determined by: (i) their relative size, (ii) how different their demand curve is from the other investors, and (iii) how price sensitive the other investors are, that is, how much do prices need to move for other investors to absorb the demand, as so far as it deviated from the market portfolio.

To compute the counterfactual using the estimated demand system, we set  $\beta_{0i} = 1$ ,  $\beta_{1i} = e_1\sigma_b$ , and  $\epsilon_i(n) = 0$ , for all  $n$ . We do not adjust the extensive margin. We then compute asset prices using this demand system by ensuring that asset markets clear. Appendix B provides the computational details. We can then explore how much asset prices change on average, but we can also re-run the valuation regressions as in Section III to explore how much different investors matter to incorporate information about fundamentals into prices.

### *C. How much do different investors matter for asset prices?*

To measure the average change in valuations, we compute the following statistic

$$\theta = \frac{1}{T} \sum_t \left( \frac{\sum_{n=1}^N |ME_t^{CF}(n) - ME_t(n)|}{\sum_{n=1}^N ME_t(n)} \right), \quad (33)$$

which measures the total repricing if we change a group's demand to holding the market portfolio. We report the results in Table 7, in which the top panel reports the results

for asset prices in the US and the bottom panel for asset prices in GB. In computing the counterfactual, we switch, for instance, hedge funds both in the US and GB to the market portfolio and solve for asset prices. In doing so, we allow investors also to rebalance across countries according to (21)-(22).

In the first column, we report  $\theta$ , the total repricing in each counterfactual scenario. In the second column, we report the fraction of all institutional capital managed by a particular group, and in the final column the fraction of all assets, which includes institutions and the household sector.

In both countries, there is a strong relation between the size of the sector and the impact on prices, which is to some extent mechanical. We summarize this graphically in the top panel of Figure 4 for the US. The first (green) bar for each sector is the fraction of total assets managed and the second (orange) bar is the change in valuations.

However, the impact is far from proportional. The bottom panel of Figure 4 displays the ratio of repricing to the ownership share for each institutional type. The ratio is close to 0.5 for mutual funds and long-term investors, approximately 0.8 for private banking and broker dealers,<sup>12</sup> and over 1.5 for hedge funds. This points to the relative importance of hedge funds in determining asset prices, while long-term investors are closer to holding a market-weighted portfolio. These numbers are particularly relevant in the context of the debate regarding the shift from active to passive investing (see for instance [Garleanu and Pedersen \(2019\)](#)).

Across countries, we find that switching investors to the market portfolio has a larger impact in the US compared to GB. The effects in GB are smaller for two reasons: (i) the household sector is larger and (ii) our estimates imply that the household sector in GB is more price elastic than the household sector in the US. Any shock to demand curves of institutions therefore has a smaller impact on asset prices.

In Table 8, we list the firms most affected by the changes in the demand system in the US. We find that the top firms affected differ substantially across institutional types, which once more illustrates the heterogeneity across investors in terms of their preferences for

---

<sup>12</sup>We only measure the direct impact of changing the holdings of the broker-dealer sector. It may well be the case, however, that by restricting leverage, broker-dealers have an outsized effect on other investors, which has a larger overall impact on prices.

characteristics and latent demand. In traditional asset pricing models, in which demand is highly elastic, such differences are largely irrelevant. However, we estimate these effects to be substantial. For instance, Apple’s value would be significantly higher if investment advisors would implement a passive indexing strategy, while Alphabet (Google) would benefit the most if hedge funds would switch to investing in the market portfolio.

#### *D. Linking characteristics to valuation ratios and long-horizon expected returns*

It is common practice in empirical asset pricing to uncover characteristics that are linked to either valuations or expected returns and to provide narratives or formal models based on investor types, such as retail investors, smart money (e.g., hedge funds), or long-term investors such as pension funds and insurance companies. Our framework can be used to understand how important different groups are to connect characteristics to prices and expected returns.

To this end, we re-run the valuation regressions by regressing the counterfactual log market-to-book ratios on characteristics in a panel with time fixed effects, as in (17). We compute the change in loadings on characteristics (multiplied by 100).

The results are presented in Table 9 for the US (top panel) and GB (bottom panel). The columns correspond to different counterfactuals. To interpret the coefficients, the estimate of 8.11 on dividend-to-book equity for investment advisors means that the valuation difference of firms with a one standard deviation difference in dividend-to-book equity would increase by 8.11% if investment advisors switch to strict market weights. We also report the change in the R-squared value in the bottom line of each panel.

We find the largest change in the coefficients for investment advisors, which is primarily driven by their size, as discussed before. However, hedge funds still play an important role in incorporating information about firm fundamentals into asset prices, which is particularly remarkable given the smaller amount of assets that they manage compared to investment advisors.

To map changes in valuations, and their connection to characteristics, to expected returns, we use the valuation model of [Cohen, Polk, and Vuolteenaho \(2003\)](#) and [Campbell, Polk,](#)

and Vuolteenaho (2009). We write the log market-to-book ratio of firm  $n$ ,  $mb_t(n)$ , as

$$mb_t(n) = \sum_{s=1}^{\infty} \rho^{s-1} \mathbb{E}_t [e_{t+s}(n)] - \sum_{s=1}^{\infty} \rho^{s-1} \mathbb{E}_t [r_{t+s}(n)], \quad (34)$$

where

$$e_t(n) = \ln \left( 1 + \frac{\Delta BE_t(n) + D_t(n)}{BE_{t-1}(n)} \right), \quad (35)$$

$$r_t(n) = \ln \left( 1 + \frac{\Delta ME_t(n) + D_t(n)}{ME_{t-1}(n)} \right), \quad (36)$$

and  $BE_t(n)$  a firm's book equity,  $ME_t(n)$  its market equity, and  $D_t(n)$  its dividend.<sup>13</sup>

To convert these estimates to expected returns, we make the simplifying assumption that expected growth rates,  $g_t$ , and expected returns,  $\mu_t$ , are random walks, which is not unreasonable given the extreme persistence in these series. The expression for the market-to-book ratio now simplifies to

$$mb_t(n) = C + \frac{g_t}{1 - \rho} - \frac{\mu_t}{1 - \rho}.$$

If the link between characteristics and expected growth rates does not change in the counterfactuals, then the change in valuation ratios links one-to-one to changes in expected returns, with a scaling coefficient of  $(1 - \rho)^{-1}$ . Using a typical value of  $\rho = 0.95$ , we obtain that the scaling factor is around 20 in mapping changes in valuations to changes in expected returns. The impact on expected returns would be larger in case expected returns are persistent but not a random walk.<sup>14</sup>

Hence, using the estimate of 8.11 as before, the relation between dividend-to-book equity and expected returns would change by 41bp per year for a one standard deviation change in dividend to book equity. If expected returns are less (more) persistent, for instance because characteristics are less (more) persistent, then these effects would be larger (smaller).

---

<sup>13</sup>As we use characteristics throughout this paper, Appendix C shows how one could compute variance decompositions in characteristics space.

<sup>14</sup>As a point of reference, the scaling coefficient equals  $(1 - \rho\varphi_\mu)^{-1}$  if expected returns follow an AR(1) with autoregressive parameter  $\varphi_\mu$ . Using the estimates in Binsbergen and Kojien (2010), the scaling coefficient is  $(1 - 0.932 \times 0.969)^{-1} \simeq 10$ .

## VII. Conclusion

It is common practice to decompose levels and variation in prices into expected returns and expected fundamentals. However, it is unclear what information investors use for prices to be informative and how important different investors are for incorporating information into prices.

We show that a small set of characteristics explains the majority of the variation in a panel of firm-level valuation ratios across countries. The same characteristics also predict future profitability with comparable coefficients. To measure how investors' demands respond to the characteristics, we estimate a demand system in Great Britain and in the United States. The demand system allows us to quantify the importance of different institutional types (e.g., mutual funds, broker dealers, ...) for price formation by computing counterfactual prices if a particular type were to follow a passive investment strategy. By combining these estimates with our forecasts of future profitability, we measure the contribution of each institutional type to cross-sectional variation in long-term expected returns.

Our framework can be used whenever one is interested in understanding why certain characteristics affect the cross-section of valuation ratios or long-term expected returns. For instance, our approach may be useful in understanding which investors matter most in connecting asset prices to corporate governance, or ESG factors model broadly ([Baker et al. 2018](#)), how risk is priced ([Pflueger, Siriwardane, and Sunderam 2018](#)), et cetera.

By focusing on groups of intermediaries, which may differ in terms of regulations, their funding structure, and investment horizon, our approach may inform the growing theoretical literature on intermediary asset pricing to develop micro-foundations for the demand curves that we estimate for different intermediaries.

## References

- Adrian, Tobias, Erkkko Etula, and Tyler Muir. 2014. “Financial Intermediaries and the Cross-section of Asset Returns.” *Journal of Finance* 69 (6):2557–2596.
- Asness, Clifford S., Tobias J. Moskowitz, and Lasse Heje Pedersen. 2013. “Value and Momentum Everywhere.” *Journal of Finance* 68 (3):929–986.
- Azar, Jose, Martin Schmalz, and Isabel Tecu. 2018. “Anticompetitive Effects of Common Ownership.” *Journal of Finance* 73 (4):1513–1565.
- Baker, Malcolm, Daniel Bergstresser, George Serafeim, and Jeffrey Wurgler. 2018. “Financing the Response to Climate Change: The Pricing and Ownership of U.S. Green Bonds.” *Working paper, Harvard Business School* .
- Bansal, Ravi and Amir Yaron. 2004. “Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles.” *Journal of Finance* 59 (4):1481–1509.
- Barberis, Nicholas, Andrei Shleifer, and Robert Vishny. 1998. “A Model of Investor Sentiment.” *Journal of Financial Economics* 49:307–343.
- Basak, Suleyman and Anna Pavlova. 2013. “Asset Prices and Institutional Investors.” *American Economic Review* 103 (5):1728–1758.
- Bena, Jan, Miguel A. Ferreira, Pedro Matos, and Pedro Pires. 2017. “Are Foreign Investors Locusts? The Long-Term Effects of Foreign Institutional Ownership.” *Journal of Financial Economics* 126 (1):122–146.
- Binsbergen, Jules H. van and Ralph S.J. Koijen. 2010. “Predictive Regressions: A Present-Value Approach.” *Journal of Finance* 65:1439–1471.
- Black, Fischer and Robert Litterman. 1991. “Asset Allocation Combining Investor Views with Market Equilibrium.” *Journal of Fixed Income* 1 (2):7–18.
- Brunnermeier, Markus K. and Lasse Heje Pedersen. 2009. “Market Liquidity and Funding Liquidity.” *Review of Financial Studies* 22 (6):2201–2238.

- Campbell, John Y. and John H. Cochrane. 1999. “By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior.” *Journal of Political Economy* 107 (2):205–251.
- Campbell, John Y., Christopher Polk, and Tuomo Vuolteenaho. 2009. “Growth or Glamour? Fundamentals and Systematic Risk in Stock Returns.” *Review of Financial Studies* 23 (1):305–344.
- Campbell, John Y. and Robert J. Shiller. 1988. “The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors.” *Review of Financial Studies* 1 (3):195–228.
- Chan, Yeung Lewis and Leonid Kogan. 2002. “Catching Up with the Joneses: Heterogeneous Preferences and the Dynamics of Asset Prices.” *Journal of Political Economy* 110 (6):1255–1285.
- Cohen, Randolph B., Christopher Polk, and Tuomo Vuolteenaho. 2003. “The Value Spread.” *Journal of Finance* 58 (2):609–641.
- Daniel, Kent and Sheridan Titman. 2006. “Market Reactions to Tangible and Intangible Information.” *Journal of Finance* 61 (4):1605–1643.
- Ellul, Andrew, Chotibhak Jotikasthira, and Christian T. Lundblad. 2011. “Regulatory Pressure and Fire Sales in the Corporate Bond Market.” *Journal of Financial Economics* 101 (3):596–620.
- Fama, Eugene F. and Kenneth R. French. 1992. “The Cross-Section of Expected Stock Returns.” *Journal of Finance* 47 (2):427–465.
- . 1995. “Size and Book-to-Market Factors in Earnings and Returns.” *Journal of Finance* 50 (1):131–155.
- Gabaix, Xavier. 2011. “The Granular Origins of Aggregate Fluctuations.” *Econometrica* 79 (3):733–772.
- Garleanu, Nicolae and Lasse Pedersen. 2019. “Active and Passive Investing.” *Working paper, University of California, Berkeley - Haas School of Business* .

- Guitierrez, German and Thomas Philippon. 2017. “Declining Competition and Investment in the U.S.” Working paper, NYU Stern.
- He, Zhiguo, Bryan Kelly, and Asaf Manela. 2017. “Intermediary Asset Pricing: New Evidence from Many Asset Classes.” *Journal of Financial Economics* 126 (1):1–35.
- Hoerl, Arthur E. and Robert W. Kennard. 1970. “Ridge Regression: Biased Estimation for Nonorthogonal Problems.” *Technometrics* 12 (1):55–67.
- Itzhak, Ben-David, Francesco A. Franzoni, Rabih Moussawi, and John Sedunov. 2018. “The Granular Nature of Large Institutional Investors.”
- Kacperczyk, Marcin T., Savitar Sundaresan, and Tianyu Wang. 2019. “Do Foreign Institutional Investors Improve Market Efficiency?”
- Koijen, Ralph S.J. and Motohiro Yogo. 2019. “A Demand System Approach to Asset Pricing.” *Journal of Political Economy*, Forthcoming.
- Lewellen, Jonathan. 2011. “Institutional Investors and the Limits of Arbitrage.” *Journal of Financial Economics* 102 (1):62–80.
- Lucas, Robert E., Jr. 1978. “Asset Prices in an Exchange Economy.” *Econometrica* 46 (6):1429–1445.
- Makarov, Dmitry and Astrid Schornick. 2010. “A Note on Wealth Effect under CARA Utility.” *Finance Research Letters* 7 (3):170–177.
- Melitz, Marc. 2003. “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity.” *Econometrica* 71 (6):1695–1725.
- Merton, Robert C. 1987. “A Simple Model of Capital Market Equilibrium with Incomplete Information.” *Journal of Finance* 42 (3):483–510.
- Pflueger, Carolin, Emil Siriwardane, and Adits Sunderam. 2018. “A Measure of Risk Appetite for the Macroeconomy.” *Working paper, UBC* .

- Sharpe, William F. 1964. "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk." *Journal of Finance* 19 (3):425–442.
- Shiller, Robert J. 1981. "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?" *American Economic Review* 71 (3):421–436.
- Vuolteenaho, Tuomo. 2002. "What Drives Firm-Level Stock Returns?" *Journal of Finance* 57 (1):233–264.

Table 1  
Firm-Level Fundamentals Granularity

Region	Mkt Pct	Size Rank	Sales Frac	NI. Frac	Inv. Frac	Emp. Frac
<b>US (2016)</b>						
	10	5	6.4	14.8	7.9	2.0
	20	14	13.1	26.7	15.3	6.9
	30	26	22.2	40.3	22.2	18.0
	40	47	29.4	52.6	25.4	26.2
	50	82	38.1	63.4	32.5	33.0
	60	138	50.2	75.4	45.9	42.1
	70	231	63.7	84.1	60.0	52.5
	80	396	72.8	91.4	71.8	61.8
	90	761	84.6	99.3	84.3	77.3
	100	3622	100.0	100.0	100.0	100.0
<b>US (2006)</b>						
	10	6	7.7	14.7	7.4	3.2
	20	14	15.9	24.7	13.7	12.1
	30	27	24.3	36.4	23.8	16.8
	40	49	31.2	46.0	29.1	22.3
	50	84	41.2	57.7	36.7	31.3
	60	146	52.1	67.1	45.7	40.2
	70	242	63.7	76.7	58.7	51.6
	80	423	74.5	86.0	69.4	62.8
	90	872	86.5	94.6	83.0	78.3
	100	4673	100.0	100.0	100.0	100.0
<b>2016</b>						
EA	50	44	40.4	53.0	48.4	33.3
EA	90	274	87.2	90.3	89.5	81.0
JP	50	85	37.9	54.3	50.1	34.0
JP	90	682	83.3	89.9	90.4	79.7
GB	50	22	40.2	40.1	46.5	27.5
GB	90	182	87.0	90.4	86.1	79.2
<b>2006</b>						
EA	50	51	48.5	58.6	51.9	35.3
EA	90	304	87.2	92.4	88.5	80.0
JP	50	70	38.1	54.4	47.8	34.6
JP	90	657	81.5	91.5	84.4	79.6
GB	50	17	45.5	46.2	47.9	28.0
GB	90	155	89.0	92.3	88.6	80.8

Each row represents the number of firms as well as the fraction of sales, net income, investment, and employment represented by the top deciles of market cap. Firm-level fundamentals are annual from FactSet from 2006 to 2016.

Table 2  
Top firms in 2016

Name	ME	Name	ME
<b>US (2016)</b>		<b>GB (2016)</b>	
Apple Inc	609	HSBC Holdings Plc	161
Alphabet Inc	548	BP Plc	122
Microsoft Corp	480	British American Tobacco plc	106
Berkshire Hathaway Inc	402	GlaxoSmithKline Plc	94
Exxon Mobil Corp	374	AstraZeneca Plc	69
Amazoncom Inc	358	Vodafone Group Plc	66
Facebook Inc	333	Diageo Plc	66
Johnson Johnson	312	Reckitt Benckiser Group Plc	60
JPMorgan Chase Co	307	Lloyds Banking Group Plc	55
Wells Fargo Co	276	Prudential Plc	52
Name	ME	Name	ME
<b>EA (2016)</b>		<b>JP (2016)</b>	
Royal Dutch Shell Plc	237	Toyota Motor Corp	176
AnheuserBusch InBev SA/NV	205	Nippon Telegraph Telephone Corp	86
Total SA	124	NTT DoCoMo Inc	85
Unilever NV	117	Mitsubishi UFJ Financial Group Inc	83
Industria de Diseco Textil SA	107	SoftBank Group Corp	73
Siemens AG	105	KDDI Corp	62
SAP SE	105	Japan Tobacco Inc	59
L'Oral SA	102	Honda Motor Co Ltd	53
Sanofi	102	Sumitomo Mitsui Financial Group Inc	52
Medtronic Plc	98	Mizuho Financial Group Inc	46

Top 10 firms by market cap within each region in 2016. Market cap is in billions of USD. Price data is from FactSet.

Table 3  
Top Investors by Type

Type	Largest investor by type	AUM
Households		6,588
Mutual Funds	The Vanguard Group, Inc.	1,598
Investment Advisors	BlackRock Fund Advisors	1,069
Long-Term	Norges Bank Investment Management	199
Private Banking	Goldman Sachs & Co. LLC (Private Banking)	99
Hedge Funds	AQR Capital Management LLC	62
Brokers	Credit Suisse Securities (USA) LLC (Broker)	60

Largest US investors by assets under management for each type in 2016. Investor types are: Investment Advisors (IA), Mutual Funds (MF), Long-term (LT), Private Banking (PB), Brokers (BR), and Households (HH). The household sector is constructed so that total holdings of a company add up to the company's market capitalization. Equity holdings data are from FactSet.

Table 4  
Top Investors.

Name	Type	AUM
<b>US (2016)</b>		
The Vanguard Group, Inc.	MF	1598
BlackRock Fund Advisors	IA	1069
SSgA Funds Management, Inc.	MF	954
Fidelity Management & Research Co.	IA	554
T. Rowe Price Associates, Inc.	MF	423
Capital Research & Management Co. (World Investors)	IA	328
Wellington Management Co. LLP	MF	324
Northern Trust Investments, Inc.	IA	276
Capital Research & Management Co. (Global Investors)	IA	250
Norges Bank Investment Management	LT	199
<b>Total</b>		<b>5974</b>
<b>GB (2016)</b>		
The Vanguard Group, Inc.	MF	61
Legal & General Investment Management Ltd.	IA	47
Norges Bank Investment Management	LT	45
BlackRock Investment Management (UK) Ltd.	IA	43
BlackRock Advisors (UK) Ltd.	IA	28
Invesco Asset Management Ltd.	IA	28
Capital Research & Management Co. (World Investors)	IA	28
Capital Research & Management Co. (Global Investors)	IA	27
SSgA Funds Management, Inc.	MF	27
BlackRock Fund Advisors	IA	27
<b>Total</b>		<b>359</b>

Top 10 institutional investors by assets under management in the US and GB in 2016. Investor types are: Investment Advisors (IA), Mutual Funds (MF), Long-term (LT), Private Banking (PB), and Brokers (BR). Equity holdings data are from FactSet.

Table 5  
Valuation and Earnings Regressions (4 Regions)

	$mb$ (US)	$e^5$ (US)	$mb$ (EA)	$e^5$ (EA)	$mb$ (JP)	$e^5$ (JP)	$mb$ (GB)	$e^5$ (GB)
Foreign Sales	0.16 (21.60)	0.03 (1.41)	0.12 (7.74)	0.09 (6.85)	0.10 (7.73)	0.00 (0.33)	0.12 (5.86)	0.07 (1.20)
Lerner	0.06 (5.27)	0.09 (6.94)	0.11 (6.21)	0.09 (7.79)	0.13 (11.91)	0.10 (6.81)	0.06 (2.69)	0.15 (3.04)
Sales to Book	0.19 (30.50)	0.14 (8.25)	0.14 (16.64)	0.05 (3.61)	0.15 (17.53)	0.10 (20.08)	0.21 (5.84)	0.20 (2.26)
Dividend to Book	0.16 (16.27)	0.15 (8.47)	0.20 (14.36)	0.19 (6.08)	0.19 (17.07)	0.04 (1.92)	0.32 (11.58)	0.23 (4.36)
Market Beta	-0.06 (-3.19)	-0.00 (-0.02)	-0.04 (-2.63)	-0.04 (-2.15)	-0.01 (-0.31)	0.02 (1.12)	-0.04 (-1.75)	0.02 (0.83)
LNbe	-0.46 (-36.10)	-0.19 (-8.48)	-0.43 (-47.88)	-0.24 (-30.37)	-0.23 (-12.24)	-0.11 (-7.13)	-0.45 (-12.73)	-0.28 (-7.76)
Adj. R <sup>2</sup>	0.54	0.29	0.61	0.36	0.42	0.38	0.70	0.54
Within Adj. R <sup>2</sup>	0.52	0.27	0.56	0.33	0.37	0.17	0.68	0.51
Num. obs.	8537	3100	3027	1124	7100	2800	1641	544

Regressions of valuation ratios on firm-level characteristics for 4 regions: United States (US), Euro Area (EA), Japan (JP), and Great Britain (GB). All regressions include year fixed effects.  $mb$  is the log market-to-book ratio at time  $t$ .  $e^5$  is cumulative return on equity from time  $t$  to  $t + 5$  adjusted for repurchases. Characteristics are measured at time  $t$ . Foreign sales is the fraction of sales from abroad, and Lerner is operating income after depreciation divided by sales, market beta is 60-month rolling market beta where the market is the local MSCI index, and  $LNbe$  is log book equity. Firm-level fundamentals are from FactSet from 2006 until 2016.

Table 6  
Demand Estimation Summary

	Investment Advisors	Households	Mutual Funds	Hedge Funds	Long-Term	Private Banking	Brokers
<b>US</b>							
LNmebe	0.746	0.019	0.917	0.256	0.818	0.711	0.924
Foreign Sales	0.006	-0.045	0.034	-0.109	0.003	0.049	0.024
Lerner	0.016	0.027	0.012	-0.038	0.025	0.044	-0.017
Sales to Book	-0.004	0.025	0.021	-0.060	0.026	0.027	0.067
Dividend to Book	-0.087	0.212	-0.091	-0.177	-0.006	0.008	0.052
Market Beta	0.020	-0.098	0.006	0.044	0.000	-0.066	0.055
LNbe	-0.129	0.289	-0.016	-0.508	0.014	-0.099	0.167
<b>GB</b>							
LNmebe	0.583	-1.905	0.276	-0.106	0.755	0.650	0.779
Foreign Sales	0.004	0.012	0.094	-0.075	0.017	-0.104	-0.081
Lerner	-0.008	0.040	-0.011	0.056	0.062	-0.039	0.005
Sales to Book	-0.026	0.000	0.063	-0.006	0.048	-0.030	0.017
Dividend to Book	0.039	-0.016	0.009	-0.265	-0.061	-0.010	-0.067
Market Beta	-0.103	0.069	-0.080	-0.065	-0.077	-0.094	-0.039
LNbe	-0.089	0.110	0.009	-0.486	-0.004	-0.191	-0.109
<b>International</b>							
Cross-country	0.317		0.167	0.126	0.267	0.214	0.098

Summary statistics of coefficient estimates from investor level demand system estimation in the United States (US) and Great Britain (GB). AUM is the time series average of the total AUM for each investor group. All other cells are the time series average of the within year assets under management weighted average coefficients. *LNmebe* is the coefficient on log market-to-book ratio. The remaining coefficients are the deviations from the market valuation regression coefficients. Investor types are: Investment Advisors (IA), Mutual Funds (MF), Long-term (LT), Private Banking (PB), Brokers (BR), and Households (HH). The household sector is constructed so that total holdings of a company add up to the company's market capitalization. Equity holdings data are from FactSet from 2006 until 2016.

Table 7  
Counterfactual market cap change.

	Change	frac Inst	frac Tot
<b>US</b>			
Investment Advisors	30.2	56.3	39.4
Mutual Funds	12.0	27.1	19.0
Hedge Funds	6.2	5.3	3.7
Long-Term	1.8	5.8	4.0
Private Banking	2.5	3.9	2.7
Brokers	1.0	1.6	1.1
<b>GB</b>			
Investment Advisors	8.6	68.5	40.0
Mutual Funds	2.2	13.1	7.7
Hedge Funds	0.4	1.2	0.7
Long-Term	0.9	9.4	5.5
Private Banking	0.7	3.6	2.1
Brokers	0.8	4.3	2.5

Each change is the time series average of the total absolute change in firm's market caps under counterfactual market equities divided by the total original market cap. frac Inst is the time series average of the fraction of institutional ownership by that type and frac Tot is the fraction of total ownership. The counterfactual market equities are calculated under the assumption that each investor type follows strict market weights. Investor types are: Investment Advisors (IA), Mutual Funds (MF), Long-term (LT), Private Banking (PB), and Brokers (BR). Firm-level fundamentals and equity holdings data are from FactSet from 2006 until 2016.

Table 8  
Largest Changes in Market Cap (US)

Company	ME	ME (CF)
<b>IA</b>		
Apple, Inc.	608.7	1025.0
Alphabet, Inc.	547.8	864.4
UDR, Inc.	9.7	0.9
CubeSmart	4.8	0.4
<b>MF</b>		
DDR Corp.	5.6	6.9
Berkshire Hathaway, Inc.	402.0	495.3
DexCom, Inc.	5.1	1.8
ServiceNow, Inc.	12.4	4.1
<b>HF</b>		
Alphabet, Inc.	547.8	655.7
Microsoft Corp.	480.3	565.0
Seattle Genetics, Inc.	7.5	1.8
Icahn Enterprises LP	8.7	1.5
<b>LT</b>		
HCP, Inc.	13.9	15.8
American Campus Communities, Inc.	6.6	7.4
Macerich Co.	10.2	5.0
CNA Financial Corp.	11.2	2.4
<b>PB</b>		
Hudson Pacific Properties, Inc.	4.7	11.3
Targa Resources Corp.	10.4	12.0
Norwegian Cruise Line Holdings Ltd.	9.7	5.7
Hilton Worldwide Holdings, Inc.	26.9	10.5
<b>BR</b>		
American Campus Communities, Inc.	6.6	6.8
CarMax, Inc.	12.1	12.4
Citrix Systems, Inc.	14.0	11.9
CONSOL Energy, Inc.	4.2	3.5

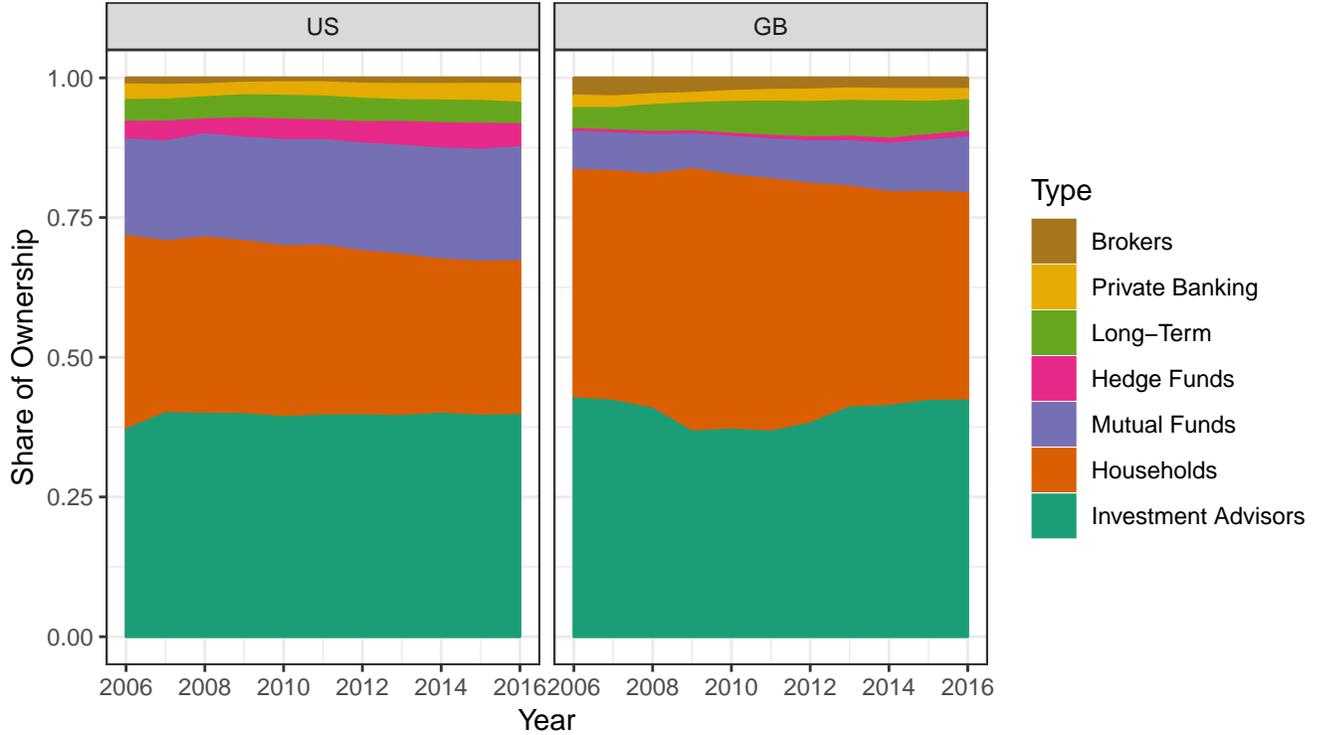
Two largest percent increases and decreases in market cap in 2016 under counterfactual market equities. The counterfactual market equities are calculated under the assumption that each investor type follows strict market weights. Investor types are: Investment Advisors (IA), Mutual Funds (MF), Long-term (LT), Private Banking (PB), and Brokers (BR). Firm-level fundamentals and equity holdings data are from FactSet from 2006 until 2016.

Table 9  
Change in Counterfactual Valuation Regression Coefficients.

	IA	MF	HF	LT	PB	BR
<b>US</b>						
Foreign Sales	-0.74	-1.38	1.41	0.13	-0.19	0.10
Lerner	1.22	0.55	1.36	-0.18	-0.37	0.07
Sales to Book	5.35	0.35	-1.05	0.23	-0.23	-0.03
Dividend to Book	8.11	3.09	3.65	-0.28	0.19	-0.11
Market Beta	-1.92	0.20	-2.11	-0.38	0.02	-0.24
LNbe	21.44	4.02	4.87	0.07	1.12	-0.07
R-squared	-21.71	-3.99	-5.11	-1.34	-1.20	-0.18
<b>GB</b>						
Foreign Sales	0.30	-0.56	0.05	0.33	0.25	0.30
Lerner	1.19	0.55	-0.11	-0.48	0.04	0.02
Sales to Book	0.89	0.08	-0.06	-0.16	-0.08	0.11
Dividend to Book	1.27	0.17	0.12	0.60	0.10	0.10
Market Beta	1.45	0.01	0.15	0.24	0.03	-0.09
LNbe	6.67	0.88	0.33	0.19	0.27	0.40
R-squared	-6.01	-0.52	-0.22	-0.32	-0.25	-0.33

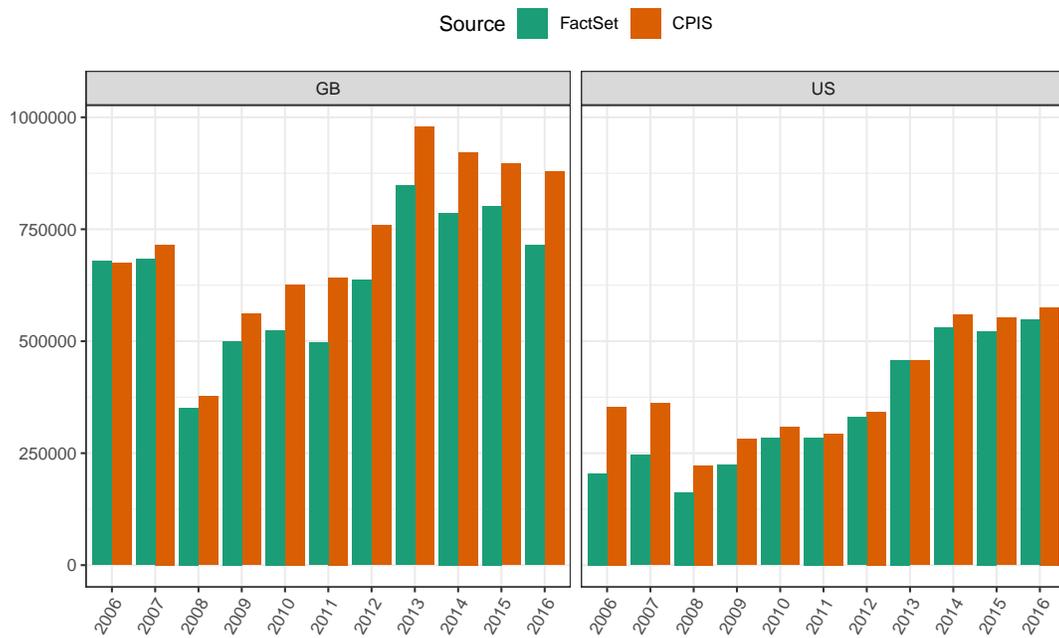
Change in counterfactual regressions of valuation ratios on firm-level characteristics. Each column is the change in the coefficient from the actual valuation regression multiplied by 100. The counterfactual market-to-book ratios are calculated under the assumption that each investor type follows strict market weights. All regressions include year fixed effects.  $mb$  is the log market-to-book ratio at time  $t$ . Characteristics are measured at time  $t$ . Foreign sales is the fraction of sales from abroad, and Lerner is operating income after depreciation divided by sales, and market beta is 60-month rolling market beta where the market is the local MSCI index, and  $LNbe$  is log book equity. Investor types are: Investment Advisors (IA), Mutual Funds (MF), Long-term (LT), Private Banking (PB), Brokers (BR), and Households (HH). Firm-level fundamentals and equity holdings data are from FactSet from 2006 until 2016.

Figure 1  
Time series of ownership by institutional type.



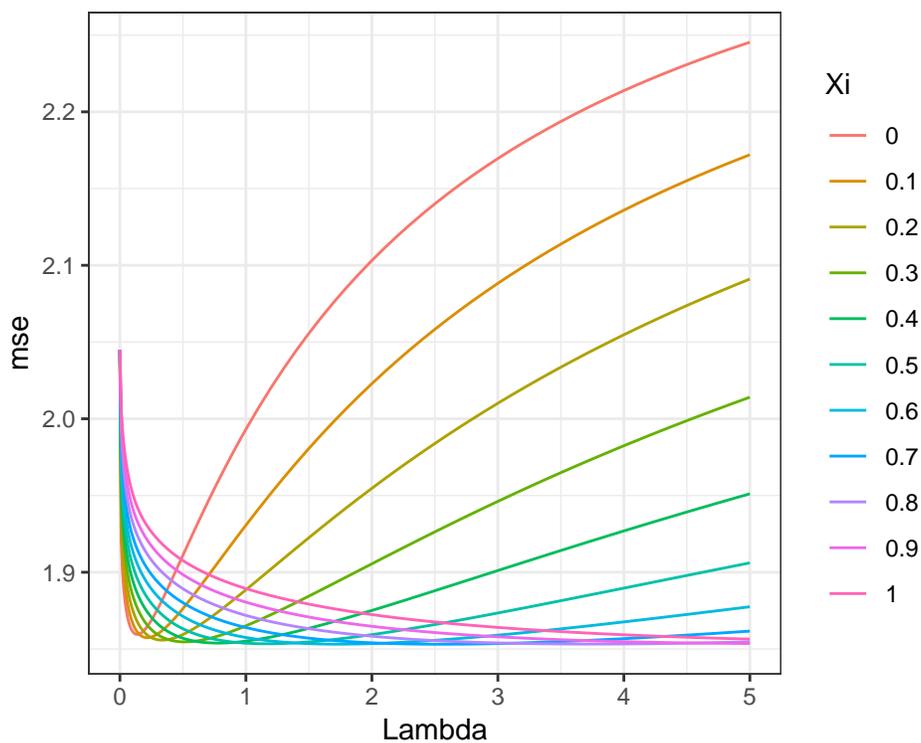
Summary statistics of coefficient estimates from investor level demand system estimation in the United States (US) and Great Britain (GB). AUM is the time series average of the total AUM for each investor group. All other cells are the time series average of the within year assets under management weighted average coefficients.  $LNmebe$  is the coefficient on log market-to-book ratio. The remaining coefficients are the deviations from the market valuation regression coefficients. Investor types are: Investment Advisors (IA), Mutual Funds (MF), Long-term (LT), Private Banking (PB), Brokers (BR), and Households (HH). The household sector is constructed so that total holdings of a company add up to the company's market capitalization. Equity holdings data are from FactSet from 2006 until 2016.

Figure 2  
Comparison with IMF CPIS.



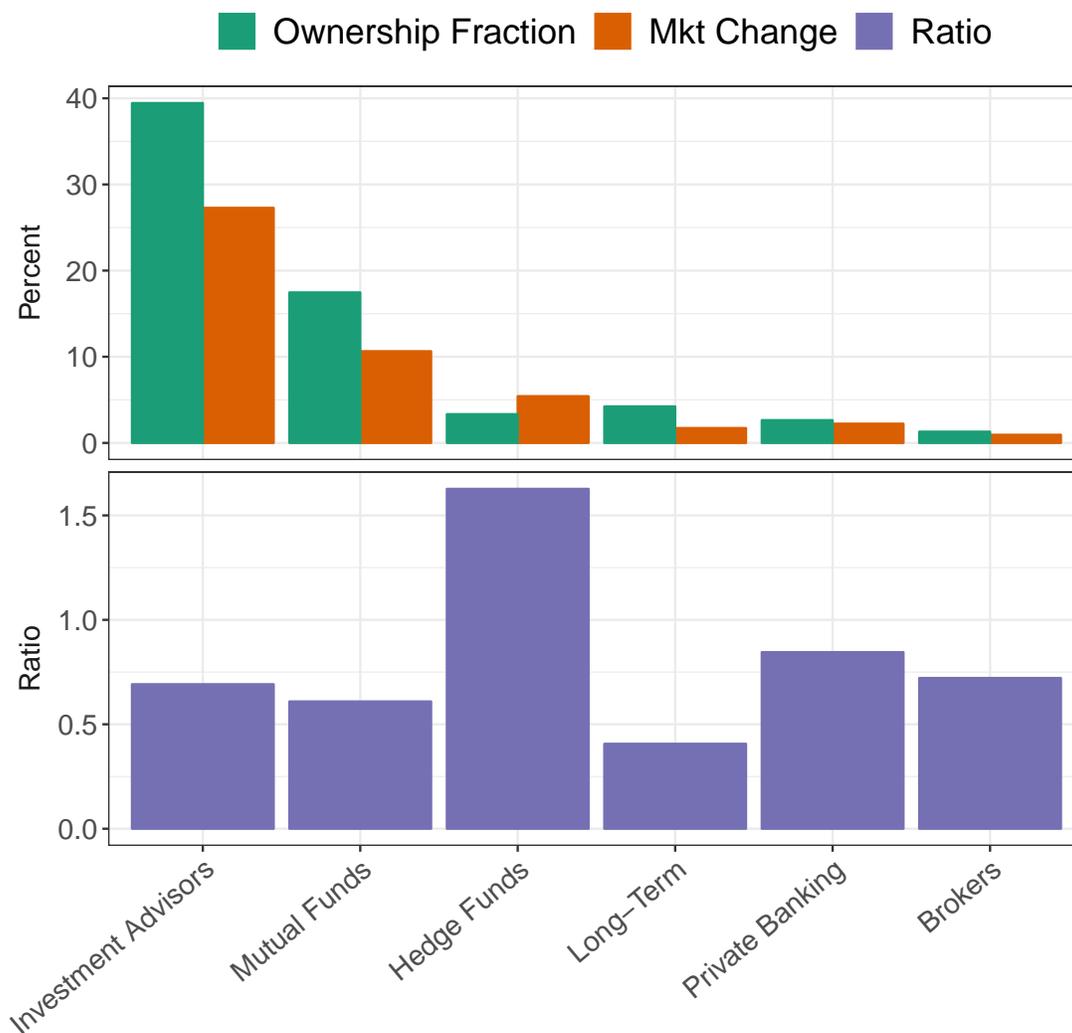
Total cross-border holdings of US and UK equities by US and UK investors. Equity holdings are from the IMF Coordinated Portfolio Investment Survey and FactSet from 2006 until 2016.

Figure 3  
Out-of-sample mean squared error in demand estimation.



Out-of-sample mean squared error from ridge estimator for various parameters. Lambda controls the shrinkage of coefficients and  $\xi$  controls the rate at which the shrinkage declines as the sample size increases. For each investor-year the sample is split randomly into 2 equal parts. Demand is estimated on one half and mean squared error is estimated on the other half. This figure plots the average MSE across all investors.

Figure 4  
Counterfactual of Change in Market Capitalization



Change it total market cap changing each investor type to a pure indexing strategy. The top panel reports the fraction of ownership and the total change in market cap under each counterfactual price scenario. The bottom panel reports the change in market cap normalized by the fraction of ownership by each group. Investor types are: Investment Advisors (IA), Mutual Funds (MF), Long-term (LT), Private Banking (PB), Brokers (BR), and Households (HH). Firm-level fundamentals and equity holdings data are from FactSet from 2006 until 2016.

## A. Data construction

All FactSet data are from WRDS. We use FactSet fundamentals annual version 3 and FactSet ownership version 5. MSCI return indices are from DataStream. Interest rate data are from Global Financial Data.

We combine data from these sources to build an annual end-of-year panel of firm-level fundamentals and investor-level equity holdings from 2006 until 2016. Our fundamentals data covers 4 regions: United States (US), Euro Area (EA), Great Britain (GB), and Japan (JP). Our holdings data covers 2 regions: United States (US) and Great Britain (GB). EA consists of companies in Austria, Belgium, Finland, France, Germany, Ireland, Italy, Netherlands, Portugal, and Spain.

**Market capitalization** We combine monthly security prices and company-level shares outstanding from `monthly_prices_final_usd` and `monthly_prices_final_int` with point-in-time exchange rates from `fx_rates_usd`. We calculate company-level USD market cap using shares outstanding (`ff_shs_out`) times price (`price_m`) and convert to USD. Shares outstanding is the common shares outstanding at the company level. If a company has more than one share class, shares outstanding is adjusted for the relative par values of all share classes. Both prices and shares outstanding are adjusted for splits.

Some companies have many listed securities. To merge market caps with fundamentals, we select a unique primary security `fsym_id` for each company (`factset_entity_id`). We start from the set of securities which we can calculate a USD market cap for and merge on security and firm identifying information from `sym_coverage` and `ff_sec_coverage`. We sequentially select the first security for each company which is uniquely identified by the following criteria: one security for the company, `ff_iscomp` is 1 for the security, the security is identified as primary by `fsym_primary_equity_id`. If this procedure does not uniquely identify a primary security, we do not include the company in our sample. This occurs in a very small number of cases.

For each firm-year we use the market cap reported at the end of December.

**Fundamentals** Fundamentals are in 4 files for each of 3 regions: `ff_FILE_REGION` where `FILE` is one of `basic_af`, `basic_der_af`, `advanced_af`, `advanced_der` and `REGION` is one of `ap`, `eu`, `am`. We merge these 12 files together with point-in-time exchange rates from `fx_rates_usd` and convert monetary values to USD. For fundamentals in December of each year we use the most recently available fundamentals as of the end of June of that year.

- Book equity is total shareholders equity (`ff_shldrs_eq`) plus deferred taxes and investment tax credits (`ff_dfd_tax_ite`) minus preferred stock (`ff_pfd_stk`). We set preferred stock to zero if it does not exist and drop negative book equity values.
- Market equity is total value of common equity as detailed in the market capitalization section above.
- Foreign sales share is international sales (`ff_sales_intl`) divided by total sales (`ff_sales`).

- Lerner is operating income before depreciation (`ff_oper_inc_bef_dep`) minus depreciation and amortization (`ff_dep_amort_exp`) if available or operating income (`ff_oper_inc`) divided by sales (`ff_sales`).
- Sales to book is sales (`ff_sales`) divided by book equity
- Dividends to book are dividends (`ff_div_cf`) divided by book equity.
- Betas are from 60-month rolling regressions on MSCI local equity market index returns. Excess returns are calculated using 3-month rates from Global Financial data.
- Net repurchases are `ff_stk_purch_cf` minus `ff_stk_sale_cf` and are set to 0 if missing.

We Winsorize beta at the 2.5% and 97.5% level and Winsorize dividend-to-book, and sales-to-book at the 97.5% level by region-year. We set values of Lerner that are less than -1 to -1.

**Portfolio Holdings** We build a panel of end-of-year equity holdings of US and GB companies for institutional and non-institutional investors. FactSet collects data for global companies and institutions, but the coverage outside of the US and GB is not sufficient for our purposes of estimating a demand system. To remain consistent with FactSet’s methodology we construct holdings data for all countries and then select holdings of only US and GB companies. We also limit our sample to 2006 until 2016 due to lower coverage in GB prior to 2006.

FactSet holdings data are from 4 broad sources:

- 13F holdings (13F). 13F data are from mandatory 13F reports on US-traded equities held by institutions managing more than \$100 million in US-traded securities. Data is in `own_inst_13f_detail_eq`.
- Sum of fund-level reports (SOF). Fund-level data are from SEC mandated reports in the US and are collected directly from funds managers by FactSet in other countries. Data is in `own_fund_detail_eq`. Fund-level reports are aggregated to the institution level using the mapping from fund ids to institution ids in `own_ent_funds`.
- Institutional Stakes (INST). Institutional stakes data for GB and are sourced from share registers (UKSR) and regulatory news service filings (RNS). FactSet analyzes share registers at minimum annually, though for companies larger than fledgling the frequency is quarterly. Institutional stakes data for the US are sourced from regulatory filings such as 10K, 13D, 13G, and proxies. For other countries FactSet collects data from various regulatory filings. Data is in `own_inst_stakes_detail_eq`.
- Non-institutional stakes (NISTK). Non-institutional stakes are from regulatory filings and primarily represent holdings by firm insiders or by other companies. Data is in `own_stakes_detail_eq`.

We combine data from the 4 sources. Securities are identified as either 13F US (`fds_13f_flag=1`), 13F Canada (`fds_13f_ca_flag=1`), or UKSR (`fds_uksr_flag=1`) in `own_sec_coverage`. Holder's are identified as 13F institutions (`fds_13f_flag=1`) in `own_ent_institutions`. We use the following rules to combine institutional holdings as is done by FactSet:

- For UKSR securities, select UKSR or RNS positions (types W and Q) if the `as_of_date` is within 18 months of December of each year. If there are no institutional stakes based filings for a given institution use SOF if the report is within 18 months of December of each year.
- For 13F institutions and 13F US securities, use the 13F position if it is within 18 months of December of each year unless there is a more recent INST position.
- For 13F institutions and 13F CA securities, use the 13F position if it is within 18 months of December of each year unless there is a more recent INST position. If there is are no 13F or INST positions, use SOF if it is within 18 months of December of each year.
- For non-13F institutions and/or non-13F US/CA securities, use the INST position if it is within 18 months of December of each year for US securities and 21 month for non-US securities. If there is no INST position, use SOF if it is within 18 months of December of each year.
- Use NISTK positions if they are within 18 months of December of each year.

We merge on prices from `own_sec_prices_eq` and calculate dollar values of holdings for holdings of each security. We limit holdings to common equity and ADRs:

(`fref_security_type=SHARE,ADR,DR,GDR,NVDR` and `issue_type=EQ` in `sym_coverage`)

We aggregate dollar values of security-level holdings to company-level holdings using the mapping in `own_sec_entity_eq`.

We classify institutions into types using FactSet's `investor_sub_type` in `sym_entity`. Hedge Fund=AR, FH, FF, FU, FS; Broker=BM, IB, ST, MM; Private Banking=CP, FY, VC; Investment Advisor=IC, RE, PP, SB; Long-term=FO,SV,IN; Mutual Fund = MF.

We construct the Household sector so that total holdings of institutions and household are equal to each firm's market cap. On occasion, total holdings of institutions are greater than the market cap, in which case we proportionally scale back all institutions holdings.

We classify the outside asset as any firm which is outside of the top 90% of market cap within each region. Any institution which has less than \$1mm in holdings in the outside asset, is classified as a non-institutional stakes holder, or has less than 20 holdings across all years is moved into the household sector.

## B. Computing counterfactual asset prices

To compute the counterfactual asset prices in this case, we start from

$$\begin{aligned}
 y_{it}(n) &= \ln \left( \frac{w_{it}(n)}{w_{it}(0)} \right) - \gamma' x_{it}^*(n) - b_t(n) \\
 &= c_{0it} + \beta_{0i} u_t(n) + c'_{1i} x_t^*(n) + \epsilon_{it}(n),
 \end{aligned}$$

where  $u_t(n) = mb_t(n) - c_{mb,t} - \gamma'x_t^*(n)$ . Combining these identities implies

$$\begin{aligned}\ln\left(\frac{w_{it}(n)}{w_{it}(0)}\right) &= c_{0it} - c_{mb,t}\beta_{0i} + \beta_{0i}mb_t(n) + (c_{1i} + \gamma(1 - \beta_{0i}) + e_1\sigma_b)'x_t^*(n) + \epsilon_{it}(n) \\ &= c_{wi,t}(\beta_{0i}) + \beta_{0i}mb_t(n) + \beta'_{1i}x_t^*(n) + \epsilon_{it}(n),\end{aligned}$$

where  $c_{wi,t}(\beta_{0i}) = c_{0i,t} - c_{mb,t}\beta_{0i}$  and  $\beta_{1i} = c_{1i} + \gamma(1 - \beta_{0i}) + e_1\sigma_b$ .

We use this last equation as the demand curve and change the coefficients and latent demand for one investor type. In particular, under the counterfactual, we assume

$$\ln\left(\frac{w_{it}(n)}{w_{it}(0)}\right) = c_{wi,t}(1) + me_t(n) = c_{0it} - c_{mb,t} + me_t(n),$$

for one type of investors. The demand curve of all other investors remains unchanged. We then solve for asset prices using the market clearing equation

$$ME_t^{CF}(n) = \sum_i w_{it}^{CF}(n, ME_t^{CF}(n))A_{it}, \quad (37)$$

where  $w^{CF}$  are the counterfactual portfolio weights. We use the algorithm in [Kojien and Yogo \(2019\)](#) to compute counterfactual market capitalizations, which iterates on (37) until convergence.

### C. Variance decompositions using characteristics

We show how our valuation regressions and earnings predictability regressions connect to traditional variance decompositions. Starting from (34) without expectations, it holds

$$mb_t(n) = c + \sum_{s=1}^{\infty} \rho^{s-1}e_{t+s}(n) - \sum_{s=1}^{\infty} \rho^{s-1}r_{t+s}(n). \quad (38)$$

Consider a linear projection of both sides on a set of characteristics,  $x_t(n)$  as well as a time fixed effect, which yields

$$a_{mb,t} + \lambda'_{mb}x_t(n) = a_{e,t} + \lambda'_e x_t(n) - (a_{r,t} + \lambda'_r x_t(n)),$$

implying

$$a_{mb,t} = a_{e,t} - a_{r,t}, \quad (39)$$

$$\lambda_{mb} = \lambda_e - \lambda_r. \quad (40)$$

Hence, the fraction of market-to-book ratios that can be explained by characteristics,  $\text{Var}(\lambda'_{mb}x_t(n))$ , satisfies the variance decomposition

$$\text{Var}(\lambda'_{mb}x_t(n)) = \text{Cov}(\lambda'_{mb}x_t(n), \lambda'_e x_t(n)) - \text{Cov}(\lambda'_{mb}x_t(n), \lambda'_r x_t(n)),$$

and the fraction due to returns therefore equals

$$\text{Fraction due to expected returns} = \frac{\lambda'_{mb} \Sigma_x (\lambda_{mb} - \lambda_e)}{\lambda'_{mb} \Sigma_x \lambda_{mb}},$$

and the fraction due to expected growth rates

$$\text{Fraction due to expected growth rates} = \frac{\lambda'_{mb} \Sigma_x \lambda_e}{\lambda'_{mb} \Sigma_x \lambda_{mb}},$$

with  $\Sigma_x = \text{Var}(\Sigma_x)$ . As characteristics are cross-sectionally standardized, if the characteristics are also uncorrelated, then the shares equal  $\frac{\lambda'_{mb} (\lambda_{mb} - \lambda_e)}{\lambda'_{mb} \lambda_{mb}}$  and  $\frac{\lambda'_{mb} \lambda_e}{\lambda'_{mb} \lambda_{mb}}$ , respectively.