All Actuaries Summit
Think bigger, live better
1-3 May 2024 • Gold Coast

Actuaries Institute.

Scale Diseconomies and Capacity in Fund Management: Variation Across Equity Markets

Prepared by Michael J. O'Neill, Jie (Felix) Sun, Geoffrey J. Warren, Min Zhu

Presented to the Actuaries Institute
2024 All-Actuaries Summit
1-3 May 2024

Actuaries Institute
ABN 69 000 423 656
Level 2, 50 Carrington Street, Sydney NSW 2000
P +61 (0) 2 9239 6100 | actuaries.asn.au

Scale Diseconomies and Capacity in Fund Management: Variation Across Equity Markets

May 12, 2022

Michael J. O'Neill
Portfolio Manager, Investors Mutual Limited
Adjunct Associate Professor, Bond Business School
Level 24, 25 Bligh St, Sydney, NSW, AUSTRALIA, 2000
+61 402 357 582
michael.oneill@iml.com.au

Jie (Felix) Sun
Principal, Mercer Australia
100 Barangaroo Avenue, Barangaroo NSW, AUSTRALIA, 2000
+61 2 8864 6957
Felix.Sun@mercer.com

Geoffrey J. Warren
Australian National University
Level 4, CBE Building (26C), Kingsley Street, Acton, ACT, AUSTRALIA, 2601
+61 411 241 091
geoff.warren@anu.edu.au

Min Zhu
University of Queensland
Room 321, Colin Clark Building, St. Lucia Campus, Brisbane QLD, AUSTRALIA, 4072
+61 7 334 68147
m.zhu@business.uq.edu.au

JEL classification: G11, G23, J24

Keywords: Equity mutual funds, capacity constraints, economies of scale

Declaration: The authors would like to thank Mercer for providing the fund data used in this study. Mercer makes no representations or warranties as to the accuracy of the information presented and takes no responsibility or liability (including for indirect, consequential or incidental damages), for any error, omission or inaccuracy in the data supplied by investment managers. Past performance is not an indication of future performance. Contributions by Felix Sun to this paper were made in his personal capacity, and not as a representative of Mercer. The authors take full responsibility for the contents of this paper, which do not reflect Mercer's formal view or that of its employees.

1. Introduction

The existence of diseconomies of scale in fund management is an important topic in both theory and practice. In terms of theory, scale diseconomies suggest a market equilibrium where assets under management (AUM) gravitate to levels at which expected net excess return create no incentive for investors to supply additional capital. This perspective is exemplified by the model of Berk and Green (2004), denoted here as 'BG', which postulates that expected net excess returns equal zero while AUM and total management fees vary across funds in reflection of their relative skill. For industry practitioners, the idea of scale diseconomies as a fund grows in size manifests via the concept of 'capacity', being the AUM that should not be exceeded if investors are expected to benefit from active management (see Vangelisti, 2006; O'Neill, Schmidt and Warren, 2018; O'Neill and Warren, 2019). Much of research to date into scale and capacity in fund management is focused on US equity mutual funds and hedge funds. We broaden the scope of inquiry by examining how scale diseconomies and the relation between AUM and capacity vary across four equity markets – global equities, emerging markets, Australia core[1] and Australia small caps. This allows us to comment on how the relation between AUM and capacity varies across markets. We find that both industry size and fund size drive diseconomies of scale, and that the influence of these two drivers varies both across the markets examined and also over time. We uncover deviations from the equilibrium proposed by BG, that in turn vary in nature across markets.

The exact nature of the relation between active returns and AUM is unclear for a number of reasons. First, it is possible that scale economies may exist at lower AUM before any diseconomies begin to emerge as AUM increases. Second, scale diseconomies may operate not only at the fund level but also the industry level, to the extent that funds may be competing for similar opportunities and thus 'sharing capacity' (see Pástor and Stambaugh, 2012; Pástor, Stambaugh and Taylor, 2015; Harvey and Liu, 2017; Zhu 2018). This suggests that any relation between AUM and excess returns may have a component that relates to either the overall market, or the distribution of fund AUM within the market.[2] Third, BG hypothesize that no excess returns will be observed in equilibrium because either fund flows or fees will adjust until active returns are zero. Under their model, if all funds operate at the optimal scale for their investment approach, then variation in AUM would be observed but it may not be related to variation in active returns earned by investors. However, the BG model may be incomplete in a number of ways. Possibilities include agency effects that permit some funds to operate at other than optimal AUM (see Harvey and Liu, 2017; Zhu, 2018), inability of investors to identify manager skill (see Stambaugh, 2014; Song, 2020), or adjustment frictions that give rise to a relation between AUM and active returns that persists for a period until equilibrium is attained (see Foster and Warren, 2015; Yan, 2020; Barras, Gagliardini and Scaillet, 2022). For example, a negative relation between AUM and performance might be observed if it takes time for fund flows to react to performance; or if investors pursue outperforming funds beyond their capacity and funds are willing to accept more AUM than optimal. Fourth, the relation between active returns and AUM may vary across asset classes or time. The nature of the relation is an empirical issue on which we aim to shed further light through examining the relation between active fund returns and AUM across a range of equity markets.

Academic evidence on the relation between AUM and performance for equity mutual funds is mixed.[3] Chen et al. (2004), Yan (2008) and Chan et al. (2009) find that funds with lower AUM outperform those with larger AUM.

[1] Australia core equity funds invest substantially in large cap stocks, defined as the top 100 in Australia, but with some scope to invest in companies outside of the top 100.
[2] Capacity may also operate at the style or strategy level, which is not considered in this paper. For example, Hoberg, Kumar and Prabhala (2018) relate equity mutual fund performance to style-based competition. In hedge funds, Naik, Ramadorai and Stromqvist (2007) uncover diseconomies of scale at the broad sector level; while Forsberg, Gallagher and Warren (2022) provide evidence of capacity constraints being related to the total AUM pursuing similar strategies.
[3] Evidence of diseconomies of scale for other assets classes is also mixed. In hedge funds, Ammann and Moerth (2008), Fung et al. (2008), Ramadorai (2013), Yin (2016) and Gao, Haight and Yin (2018) all document that fund flows and larger AUM reduce alpha for hedge funds; while Getmansky (2012) finds a negative relation between AUM and returns across the total sample but with considerable variation across strategies. In fixed income. Phillpott et al. (1998) provide evidence of economies of scale in bond mutual funds; whereas Yan (2019) finds a hump-shaped relation between active return and AUM

On the other hand, Indro et al. (1999) find that performance increases with fund size up until the 9th decile, albeit based on limited data. This finding is consistent with both economies and diseconomies of scale, as well as the notion that not all funds operate at the optimal level. Another body of research fails to uncover any significant relation between AUM and performance, e.g. Grinblatt and Titman (1989); Gallagher and Martin (2005); Elton, Gruber and Blake (2012); Phillips, Pukthuanthong and Rau (2018). Meanwhile, a negative relation between AUM and active returns is found to exist outside of the US by Vidal-Garcia and Vidal (2021), but not by Ferreira et al. (2013). Recent research analyzes the relation between active returns for US equity mutual funds recognizing that manager skill is a missing variable that may be correlated with both AUM and performance, while allowing for scale diseconomies to stem from industry-wide as well individual fund effects. Under this approach, Pástor et al. (2015) find a significant negative relation between active returns and industry AUM but not fund-level AUM; while Zhu (2018) a significantly negative relation at both the fund and industry level after applying an estimator with increased power. Barras, Gagliardini and Scaillet (2022) find that excess returns of active US equity mutual are negatively related to fund AUM, with the strength of the relationship differing across funds and unaffected by the inclusion of industry size. Some authors have attempted to mitigate endogeneity concerns by examining fund returns after exogenous shocks, but the evidence here is also mixed. McLemore (2019) confirms the existence of diseconomies of scale for US equity mutual funds through examining fund mergers; while Reuter and Zitzewitz (2021) uncover no evidence of significant return adjustments for actively-managed US-based equity and bond mutual funds related to AUM shifts following changes in Morningstar fund ratings.

We contribute to this body of research by examining the relation between excess returns relative to a fund's benchmark index and both fund AUM and industry AUM for active equity funds operating in the four 'markets' of global equities, emerging markets, Australia core and Australia small caps. Analysis is conducted on quarterly panel data spanning over 14 years, using individual fund AUM to measure of fund size and total AUM of the funds in our sample to measure industry size. Our empirical approach builds on Pástor, Stambaugh and Taylor (2015), Harvey and Liu (2017) and Zhu (2018), who examine the relation between AUM and active returns for US equity mutual funds. In line with these authors, we allow for fund-specific impacts through applying fund fixed effects, which recognizes the difficulty of observing many potentially important control factors as well as variation in potential to generate excess returns (i.e. skill) across funds. Our proxies for industry size account for the possibility of shared capacity, again following Pástor, Stambaugh and Taylor (2015) and Harvey and Liu (2017). After fitting regression models to each market, we then use these models to estimate predicted net excess returns for each fund in each quarter by conditioning on fund size and industry size and an assumed fee. We then gauge whether funds are operating below or above capacity based on whether predicted net excess returns are positive or negative, respectively. This supports further analysis of the extent to which funds deviate from capacity, and an investigation of the dynamics of AUM adjustment.

Our results suggest that scale diseconomies exist and that the equilibrium described by BG does not strictly hold in the markets examined. We find a significant negative relation between excess returns and both fund size and industry size in all four markets. Our findings are consistent with Zhu (2018) for US equity mutual funds, for which they might be considered an out-of-sample test. The negative relation between excess return and AUM is similar across all markets except Australia small caps, which offers higher excess returns at lower AUM levels that is then more quickly eroded as AUM increases. This is consistent with small market size and hence lower liquidity being associated with greater return potential but also heightened diseconomies of scale.

---

for corporate bond funds, but no relation for funds that invest in treasuries. In unlisted asset classes, a general theme emerges of scale economies. Dyck and Pomorski (2011) find that pension plans reap scale economies from alternative assets, in particular their investments in private equity and direct real estate (also see Dyck and Pomorski, 2016). Evidence of scale economies in private equity is provided by Kaplan and Schoar (2005); although Lopez-de-Silanes, Phalippou and Gottschalg (2015) find scale diseconomies. Andonov, Eichholtz and Kok (2015) confirm the existence of scale economies in unlisted real estate. Andonov, Bauer and Cremers (2012) find that large pension funds reap scale economies in alternative assets, but that the reverse occurs in listed assets such as equities and fixed income.

An examination of the percentage of funds that are operating above capacity reveals time trends that vary across markets, and in the way they relate to average fund size and industry size. These findings provide evidence that capacity effects evolve dynamically, and that the drivers can differ across markets and over time. Identifying the percentage of funds that are significantly above and significantly below capacity reveals additional insights into the extent to which economically meaningful deviations exist from the BG equilibrium. The main finding is that there are substantially more funds operating significantly below capacity than above, especially earlier in the sample period for global equities and emerging markets, and throughout the sample period for Australia small caps. The propensity for more funds to be operating below rather than above capacity contrasts with results for US equity mutual funds where the converse is evident (see Zhu, 2018; Song, 2020; Roussanov, Ruan and Yanhao Wei, 2021; Barras, Gagliardini and Scaillet, 2022). We estimate the median time that funds remain either significantly above or significantly below capacity at between two and six quarters; but also find that adjustment dynamics vary across markets. In summary, our analysis suggests that the drivers of economies of scale and the extent of deviations from capacity can differ meaningfully across markets, suggesting caution over generalizing based on the findings from one market such as the US.

In addition to contributing to the fund management literature, our findings are useful for both investors and fund managers through enhancing understanding of how diseconomies of scale and its underlying drivers vary across markets. Knowledge of the AUM range over which an active investment approach might generate attractive returns can inform investors about when to become wary about placing further funds with active managers, as well as providing an indication to fund management organizations of the level of AUM at which they might consider closing a fund to new money. It can assist institutional asset owners such as pension funds to design portfolios that are well-suited to their size. For instance, smaller institutions might favor active management in assets where attractive active returns are available only at low levels of AUM, e.g. small cap equities. Conversely, large institutions would more sensibly focus their attention on asset classes where diseconomies of scale are less prevalent. The implication is that active management might be used differently by large institutions and smaller institutions across various asset markets.

This paper is organized as follows. Section 2 describes the data sources. Section 3 sets out the methods for estimating scale diseconomies and analyzing fund capacity. Section 4 presents the results. Section 5 concludes.

2. Data

The data for this study was supplied by Mercer from their proprietary Global Investment Manager Database (GIMD™), which contains information on more than 6,100 investment managers across 443 product categories with 29 product groups. For example, emerging markets have 24 sub-categories. Data is supplied on a discretionary basis by fund managers, with Mercer providing oversight on appropriate categorization and data cleaning.[4] Coverage is comprehensive, with GIMD™ aiming to monitor performance on all pooled products that are actively marketed to their clients. Our analysis is conducted at the product group level, which we will call fund 'categories' that are addressing 'markets'. We also refer to 'funds', noting that the funds within each category may comprise of multiple products (or strategies) made available by particular managers.

We extract monthly fund gross returns before fees, AUM and benchmark information over the period December 2002 to March 2017 for funds within five categories: international equity–global equity–core all countries; international equity–global equity–core developed; emerging markets–equity; Australia–equity; and Australia–equity small cap. We merge the two international equity categories, thus forming four groupings for analysis that we respectively denote as global equities, emerging markets, Australia core and Australia small caps. The analysis is conducted at quarterly intervals, as the AUM data is provided at this frequency. GIMD™ does not report fees, which may vary across investors depending on the institutional arrangements. We use the Mercer *Global asset manager fee survey* (Mercer, 2020) to set hurdle rates for the gross returns required to deliver breakeven excess returns after fees, as described in Section 4.4.

[4] A minimum AUM of $10 million is required for inclusion in the database.

We undertake considerable data cleaning and sample filtering. We inspect funds with significantly larger AUMs than other funds, and fix any data errors related to inconsistent units. We ensure that fund returns and AUMs are expressed in a common currency, either US dollars (US$) for global equity and emerging markets, or Australian dollars (A$) for Australia core and Australia small caps. Where multiple track records are available for a fund, we take the composite if available. We otherwise choose the representative return series with the longest history, and aggregate the AUMs. Funds with less than 8 quarters of AUM data are excluded. We accept up to two consecutive missing AUM observations, filling the missing data by using linear-log interpolation.

Filtering of the sample is largely aimed at ensuring that only active funds that address the overall market of interest are analyzed. We exclude passively managed funds after identifying them based on a tracking error of below 0.5%. We exclude funds that invest in particular industries, as identified by their benchmark. For example, funds benchmarked to indices such as S&P/ASX300 Resources, Nasdaq or Global ex-Australia Gold Miners are excluded for Australia. We winsorize return data at the percentiles of 0.5% and 99.5%.

GIMD™ is designed to provide a comprehensive product footprint, in line with the intended purpose of supporting analysis and selection of funds from all those available in a category. Nevertheless, the discretionary nature of data submission by fund managers raises the possibility of various data biases such as selection bias (including incubation bias), backfill bias and survivor bias. Concerns may arise for this particular study where these biases are correlated with fund AUM, so that the relation between active returns and AUMs becomes distorted. This is more likely to occur with respect to smaller funds as a consequence of selection and backfill bias, to the extent that funds which underperform due to being sub-scale do not enter the database. It seems unlikely that survivor bias would impact on the shape of the relation (i.e. regression coefficients) between excess returns and AUM. However, any bias in returns may impact our estimates of the percentage of funds that are operating at above versus below their predicted capacity, to the extent that the database contains better funds.

Representative benchmarks are selected to capture the market for each fund category in order to estimate excess returns. While funds are managed to a variety of benchmarks, the aim is to address the relation between AUM and returns within particular markets of interest, rather than the extent to which a manager outperforms their own benchmark. In this context, the relevant benchmarks should be widely used by investors while reflecting returns on the addressable market universe. For global equities, we benchmark funds from the core all countries category against the MSCI All Country World Index (ACWI), and those from the core developed category against the MSCI World Index. We benchmark emerging markets funds against the MSCI Emerging Markets Index, Australia core funds against the S&P/ASX300 Index, and Australia small caps against the S&P/ASX Small All Ordinaries.

3. Method

3.1. Main Variables

Our dependent variables are measures of gross fund performance relative to the representative benchmark index for the market. This setting assumes that fund managers aim to outperform their respective market index rather than returns adjusted for factor exposures, which in turn reflects the manner in which investors evaluate success or failure and allocate flows. This stance is consistent with how fund management is structured in global and Australian equity markets, where funds are more typically evaluated against broad market indices. Song (2020) provides evidence that investors fail to adjust for factor exposures when allocating flows, providing support for the relevance of framing the analysis around excess returns versus commonly-used benchmark indices.

After generating returns during quarter *t* by accumulating monthly returns for all funds and the benchmark, we estimate excess returns using equation (1):

$$XR_{i,t} = R_{i,t} - R_{B,t} \tag{1}$$

where XR_i is the excess return on fund i, R_i is the return on fund i, and R_B is the return on the benchmark index for fund i, and t represents time in quarters.

The main independent variables comprise proxies for fund size (*FS*) and industry size (*IS*). The structure we apply is that *FS* captures the relative size of individual funds versus their market, and *IS* reflects the total AUM of the actively managed funds addressing that market. Unfortunately, we are unable to observe the totality of active funds. We hence use our sample as a proxy, with *IS* formed using equation (2) and *FS* using equation (3). Our proxy for *IS* is the total sum of *AUM* for all the funds contained in the GIMD™ database category that are addressing the market, divided by the total market capitalization for that market. *FS* is calculated by dividing the *AUM* for fund *i* by the market capitalization of its own investable universe. A distinction between the overall 'market' and a fund's 'investable universe' is required for global equities, as explained below.

$$IS_t = \frac{\sum_{i=1}^{n} AUM_{i.t}}{MV_{u,t}} \tag{2}$$

$$FS_{i,t} = \frac{AUM_{i.t}}{MV_{i,t}} \tag{3}$$

where *IS* is industry size, FS_i is the size of fund *i*, *AUM* is assets under management, *n* is the number of funds in the category, MV_u is the market value of the benchmark index that represents the market, and MV_i is the market value of the benchmark index for fund *i*. *MVu* and *MVi* are the same for emerging markets, Australia core and Australia small caps. A different treatment is required for global equities as it comprises funds that address either all countries or developed markets, with the latter being a subset of the former. We define *IS* for global equities as the sum of *AUM* for all country, developed and emerging markets funds, relative to the market capitalization of the MSCI AWCI Index as the *MVu* proxy. Our proxies for MV_i are the market capitalization of the MSCI ACWI for all country funds and the MSCI World Index for developed market funds.

Our proxies for both *IS* and *FS* are imperfect, but unavoidably so. The key underlying assumption is that funds contained in the GIMD™ database are representative of the pool of actively managed funds addressing the market universe of interest. While the GIMD™ database is unlikely to capture all active funds, our *IS* proxy should suffice to tease out the relation between excess returns and *IS* if there is a constant scaling error, i.e. if the GIMD™ sample is a constant x% of all active funds. In this case, any fluctuations in *IS* over time will reflect shifts in the scope of active fund management within the market; while *FS* will form a good representation of the broad distribution of relative AUM across the active funds that are competing within the market. Scope for error will emerge, however, if the funds contained within GIMD™ vary as a portion of all active funds, i.e. x% is time-varying. Clearly there is scope for considerable noise arising from our proxies. However, this is more likely to bias against finding significant results, where an underlying relation exists.

Summary statistics for the data and excess returns are reported and discussed in section 4.1.

3.2. Models

We use panel data analysis to estimate the relation between *XR* and both *FS* and *IS*. Fund fixed effects are included, which accounts for differing potential to generate *XR* across funds. The latter may relate to variations in the underlying potential to generate *XR* as a consequence of skill, or perhaps related to style or factor exposures. The fund fixed effects help mitigate omitted variable bias, as pointed out by Pástor et al. (2015). Equation (4) describes the linear regression, while equation (5) describes the quadratic regression.

$$XR_{i,t} = \alpha_i + \beta_1 FS_{i.t-1} + \beta_3 IS_{t-1} + \varepsilon_{i,t} \tag{4}$$

$$XR_{i,t} = \alpha_i + \beta_1 FS_{i.t-1} + \beta_2 FS_{i,t-1}^2 + \beta_3 IS_{t-1} + \beta_4 IS_{i,t-1}^2 + \varepsilon_{i,t} \tag{5}$$

The fund-specific alphas α_i capture variation in the cross-sectional potential to generate *XR* across funds. The coefficients β_1 and β_3 in the above equations reflect the linear relation at fund-level and industry-level, respectively; while β_2 and β_4 are for non-linear relation at fund and industry level, respectively. The slope coefficients are common across all funds.

We chose not to apply the two-stage procedure of Zhu (2018) to remove the finite-sample bias that is caused by a contemporaneous correlation between fund AUM and fund return. In the current setting, since *FS* is formed by scaling fund AUM by the *MV* of their index, the contemporaneous correlation between this scaled *FS* measure and fund return is likely be very low. There is no bias for *IS* coefficient as there is literally no contemporaneous correlation between *IS* and *XR* for any particular fund. While the two-stage regression of Zhu (2018) is theoretically unbiased, it substantially increases estimating variance. We hence use the fixed-effects model as the most effective solution under the setup in this study.

We also investigate the extent to which funds are operating at an AUM that deviates from that where capacity is attained. This allows us to identify funds that may be operating in excess of capacity and hence could be 'over-funded', as well as those operating below capacity and hence may be 'under-funded' and hence able to accept additional AUM without resulting in a negative predicted *XR*. In order to establish where particular funds sit in relation to capacity at time *t*, we use the estimates from equation (4) or equation (5) to generate a predicted value for *XR* conditional on observed *FS* and *IS* allowing for a return hurdle (*h*) to cover fees and any required net alpha for investors. The calculations with respect to the linear model of equation (4) are described by equation (6), where the hat (^) represents regression estimates. The calculations with respect to equation (5) are comparable, except they would include quadratic terms.

$$\widehat{NXR}_{i,t} = \hat{\alpha}_i - \widehat{\beta_1} FS_{i.t-1} - \widehat{\beta_3} IS_{t-1} - h_i \tag{6}$$

where $\widehat{NXR}$ is predicted excess return net of the return hurdle; h_i is the return hurdle for fund *i*, comprising of fund fee and the required excess return for the investor; and the hats (^) represent regression estimates. An $\widehat{NXR}$>0 indicates that a fund is under capacity, while $\widehat{NXR}$<0 indicates that it is operating above capacity.

This specification aligns with the definition of threshold capacity as per Vangelisti (2006), i.e. the return required to achieve an objective. Given that we aim to evaluate $\widehat{NXR}_{i,t}$ and hence implicitly capacity against the equilibrium proposed by BG, we set the return hurdle (*h*) equal to the fund fee. Under this set-up, a value of $\widehat{NXR}_{i,t} = 0$ would indicate an AUM in accordance with the BG equilibrium, at which all value-add is captured by managers in the form of their aggregate fee.

Rejection of the BG model requires establishing that $\widehat{NXR}$ is significantly different from zero. A significantly negative value for $\widehat{NXR}_{i,t}$ would indicate that fund *i* is operating significantly above capacity and a significantly positive value that it is operating significantly below capacity. Undertaking these tests requires an estimate of the variance of $\widehat{NXR}_{i,t}$. To do so, note that equation (6) can be re-expressed as equation (7):

$$\widehat{NXR}_{i,t} = \overline{XR}_i - \widehat{\beta_1}(FS_{i.t-1} - \overline{FS}_i) - \widehat{\beta_3}(IS_{t-1} - \overline{IS}_i) - h_i \quad (7)$$

where $\overline{XR}_i$, $\overline{FS}_i$ and $\overline{IS}_i$ are the sample means for *XR*, *FS* and *IS* with respect to fund *i* or its market. The average fund return is described by equation (8):

$$\overline{XR}_i = \frac{1}{T}\sum_{t=1}^{T} XR_{i,t} = \frac{1}{T}\sum_{t=1}^{T}(\alpha_i + \beta_1 FS_{i.t-1} + \beta_3 IS_{t-1} + \varepsilon_{i,t}) = \alpha_i + \beta_1\overline{FS}_i + \beta_3\overline{IS}_i + \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{i,t} \quad (8)$$

There is no uncertainty in $\alpha_i + \beta_1\overline{FS}_i + \beta_3\overline{IS}_i$, so that the source of uncertainty comes from $\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{i,t}$. In general, we can use equation (9) to calculate the uncertainty of $\widehat{NXR}_{i,t}$:

$$Var\left(\widehat{NXR}_{i,t}\right) = \frac{\sigma_i^2}{T} + \left[FS_{i.t-1} - \overline{FS}_i\,,\quad IS_{t-1} - \overline{IS}_i\right]\begin{bmatrix}\sigma^2_{\widehat{\beta_1}} & \sigma_{\widehat{\beta_1}\widehat{\beta_3}} \\ \sigma_{\widehat{\beta_1}\widehat{\beta_3}} & \sigma^2_{\widehat{\beta_3}}\end{bmatrix}\begin{bmatrix}FS_{i.t-1} - \overline{FS}_i \\ IS_{t-1} - \overline{IS}_i\end{bmatrix} \quad (9)$$

where $\sigma_i^2 = Var(\varepsilon_{i,t})$, $\sigma^2_{\widehat{\beta_1}}$ is the variance of $\widehat{\beta_1}$, $\sigma^2_{\widehat{\beta_3}}$ is the variance of $\widehat{\beta_3}$, and $\sigma_{\widehat{\beta_1}\widehat{\beta_3}}$ is the covariance of $\widehat{\beta_1}$ and $\widehat{\beta_3}$, which can be extracted from the model fitting.

The estimate of $\widehat{NXR}_{i,t}$ reflects its 'true value' under equation (4) less a constant return hurdle, plus an estimation error of $\frac{\sum\varepsilon_{i,t}}{T}$ as the source of uncertainty. This leads to equation (10), which provides an estimate of the variance of $\widehat{NXR}_{i,t}$ for *T* observations. Note that *T* varies for each fund, such that $Var\left(\widehat{NXR}_{i,t}\right)$ will tend to be greater for funds with fewer observations.

$$Var\left(\widehat{NXR}_{i,t}\right) = \frac{Var(\varepsilon_{i,t})}{T} + \left[FS_{i.t-1} - \overline{FS}_i\,,\quad IS_{t-1} - \overline{IS}_i\right] cov\begin{pmatrix}\widehat{\beta_1} \\ \widehat{\beta_3}\end{pmatrix}\begin{bmatrix}FS_{i.t-1} - \overline{FS}_i \\ IS_{t-1} - \overline{IS}_i\end{bmatrix} \quad (10)$$

4. Results

We start by presenting some sample summary statistics, including the distribution of excess returns across the four markets. We then report the regression model estimates, and discuss what they reveal about the relation between AUM and *XR*. The distribution of $\widehat{NXR}_{i,t}$ is then presented and interpreted in terms of what it implies for the extent and persistence of deviations from capacity.

4.1. Summary Statistics

Table 1 reports summary statistics for our sample in each market, including the average *FS* and *XR* within fund quintiles formed by ranking funds by AUM at the start of each quarter. The statistics reveal we are working with unbalanced panels where the number of funds has increased over time. There is also considerable variation in sample size across the four markets. For instance, the average number of funds per quarter ranges from 41.5 for Australia small caps to 335.9 for global equities. Statistics for *IS* reveal that our fund sample comprises on average 5.87% of total market value for global equities, about 10.6% for emerging markets and Australia core, and 17.66% for Australia small cap. Average *FS* (i.e. fund AUM relative to market value of their benchmark index) is 0.0143% for global equities, 0.0619% for emerging markets, 0.1348% for Australia core and 0.4067% for Australia small caps. These estimates reveal that the AUM for global equity funds constitute a small portion of the market, while Australia small cap funds are comparatively large in size relative to their market. The standard deviation for *FS* indicate a reasonable amount of variation, while the quintile sorts suggest that *FS* is positively skewed. This variation is encouraging from the perspective of model estimation.

The typical fund in each market generates positive *XR*, with median quarterly *XR* estimated at 0.38% for global equities, 0.71% for emerging markets, 0.45% for Australia core and 1.78% for Australia small caps. In interpreting these *XR* numbers, remember it is also possible that the returns are buoyed by selection and backfill bias. Nevertheless, the *XR* estimates are in ballpark of those reported elsewhere after allowing for fees, such as Chen et al. (2010), Gallagher et al. (2017) Lieppold and Rueegg (2020) and Cao, von Reibnitz and Warren (2020). In any event, our primary concern is the relation between *XR* and AUM, rather than the average level of *XR* across the sample. In this regard, Table 1 reveals a tendency for *XR* to decrease moving from low to high AUM quintiles. This is particularly the case for Australia core and Australia small caps, although the progression is not monotonic for global equities and emerging markets. Nevertheless, the general trend is consistent with what might be expected if capacity effects were at play, bearing in mind that sorts by *FS* do not take *IS* into account.

Table 1: Summary Statistics

Table 1 presents summary statistics for the data in each of the four markets. Fund data is extracted from Mercer's Global Investment Manager Database (GIMD™). Data for industry size and index returns are based on the MSCI All Country World Index (ACWI) and/or the MSCI World Index for global equities, the MSCI Emerging Markets Index for emerging markets, the S&P/ASX300 Index for Australia core and the S&P/ASX Small All Ordinaries Index for Australia small caps.

	Global Equities			**Emerging Markets**			**Australia Core**			**Australia Small Caps**		
Number of Funds												
Total unique funds	636			332			124			78		
Mean per quarter	335.9			170.8			67.8			41.5		
Start of period	90			60			19			15		
End of period	447			247			97			58		
Size (%)	*Median*	*Mean*	*Std Dev*	*Median*	*Mean*	*Std Dev*	*Median*	*Mean*	*Std Dev*	*Median*	*Mean*	*Std Dev*
Industry	6.61%	5.87%	1.98%	9.76%	10.61%	2.37%	10.37%	10.55%	1.28%	15.62%	17.66%	5.31%
Fund	0.0024%	0.0143%	0.0413%	0.0216%	0.0619%	0.1266%	0.0746%	0.1348%	0.1896%	0.2406%	0.4067%	0.4402%
Fund Size Quintiles												
Q1 (Largest)	0.0317%	0.0598%	0.0768%	0.1489%	0.2226%	0.2145%	0.3085%	0.3817%	0.2919%	1.0523%	1.1035%	0.4399%
Q2	0.0070%	0.0076%	0.0029%	0.0465%	0.0519%	0.0205%	0.1281%	0.1465%	0.0646%	0.4415%	0.4774%	0.1796%
Q3	0.0024%	0.0026%	0.0008%	0.0197%	0.0226%	0.0109%	0.0709%	0.0838%	0.0467%	0.2378%	0.2546%	0.0919%
Q4	0.0009%	0.0010%	0.0003%	0.0075%	0.0096%	0.0065%	0.0400%	0.0458%	0.0305%	0.1268%	0.1335%	0.0476%
Q5 (Smallest)	0.0003%	0.0003%	0.0001%	0.0023%	0.0032%	0.0031%	0.0103%	0.0171%	0.0212%	0.0531%	0.0551%	0.0369%
Excess Returns vs. Index (%)	*Median*	*Mean*	*Std Dev*	*Median*	*Mean*	*Std Dev*	*Median*	*Mean*	*Std Dev*	*Median*	*Mean*	*Std Dev*
Pooled (Fund/Quarters)	0.38%	0.49%	2.83%	0.71%	0.92%	3.12%	0.45%	0.63%	2.33%	1.78%	2.04%	4.60%
Fund Size Quintiles												
Q1 (Largest)	0.34%	0.45%	2.57%	0.63%	0.84%	2.56%	0.37%	0.43%	2.23%	1.41%	1.56%	3.89%
Q2	0.43%	0.52%	2.66%	0.63%	0.86%	2.57%	0.40%	0.49%	2.22%	1.48%	1.66%	4.35%
Q3	0.38%	0.48%	2.77%	0.71%	0.90%	2.66%	0.50%	0.64%	2.19%	1.85%	2.20%	4.75%
Q4	0.40%	0.50%	3.06%	0.71%	0.86%	3.16%	0.50%	0.66%	2.41%	2.29%	2.26%	5.00%
Q5 (Smallest)	0.34%	0.50%	3.05%	0.92%	1.12%	4.26%	0.57%	0.90%	2.53%	2.13%	2.49%	4.90%

4.2. Model Estimation

Table 2 reports the regression model estimates, with the linear model output appearing in Panel A and the quadratic model output in panel B. The linear models reveal that *XR* declines with both *FS* and *IS*, consistent with the findings of Zhu (2018) for US equity mutual funds. All coefficients are significantly negative at the 1% level. The quadratic models are more difficult to interpret due to the presence of a mixture of positive and negative coefficients: we discuss our preference for the linear models below. We also caution that the magnitude of the coefficients should be interpreted with care, especially for *IS* where our measure is constructed from funds available in GIMD™ as a proxy for all active funds. Table A1 in the Appendix provides separate regression results for ACWI and developed market funds, which were combined to create the fund sample for global equities. Here the most noteworthy difference is that the coefficient of *IS* under the linear model is attenuated for ACWI and developed market categories relative to the combined global equities sample, standing at -5.12 for ACWI and -6.14 for developed markets versus -10.75 for global equities. This suggests that combining both sets of funds might better capture the underlying sensitivity of *XR* to *IS* by accounting for the overlap in the investment universe.

Table 2: Regression Estimates

Table 2 presents results from regressing fund excess returns relative to the benchmark index (*XR*) against fund size (*FS*) and industry size (*IS*), allowing for fund-specific fixed effects and hence intercepts. *IS* is estimated as the total sum of AUM for all the funds contained in the GIMD™ database that are addressing the market, divided by the total market capitalization for that market. *FS* is calculated by dividing the AUM for fund *i* by the market capitalization of its own investable market universe. Panel A reports results for regressions including linear terms only, while Panel B reports results for regressions include quadratic terms for both *FS* and *IS*. Panel C provides an indication of the data range by providing observations at the 1st and 99th percentiles for both *FS* and *IS*.

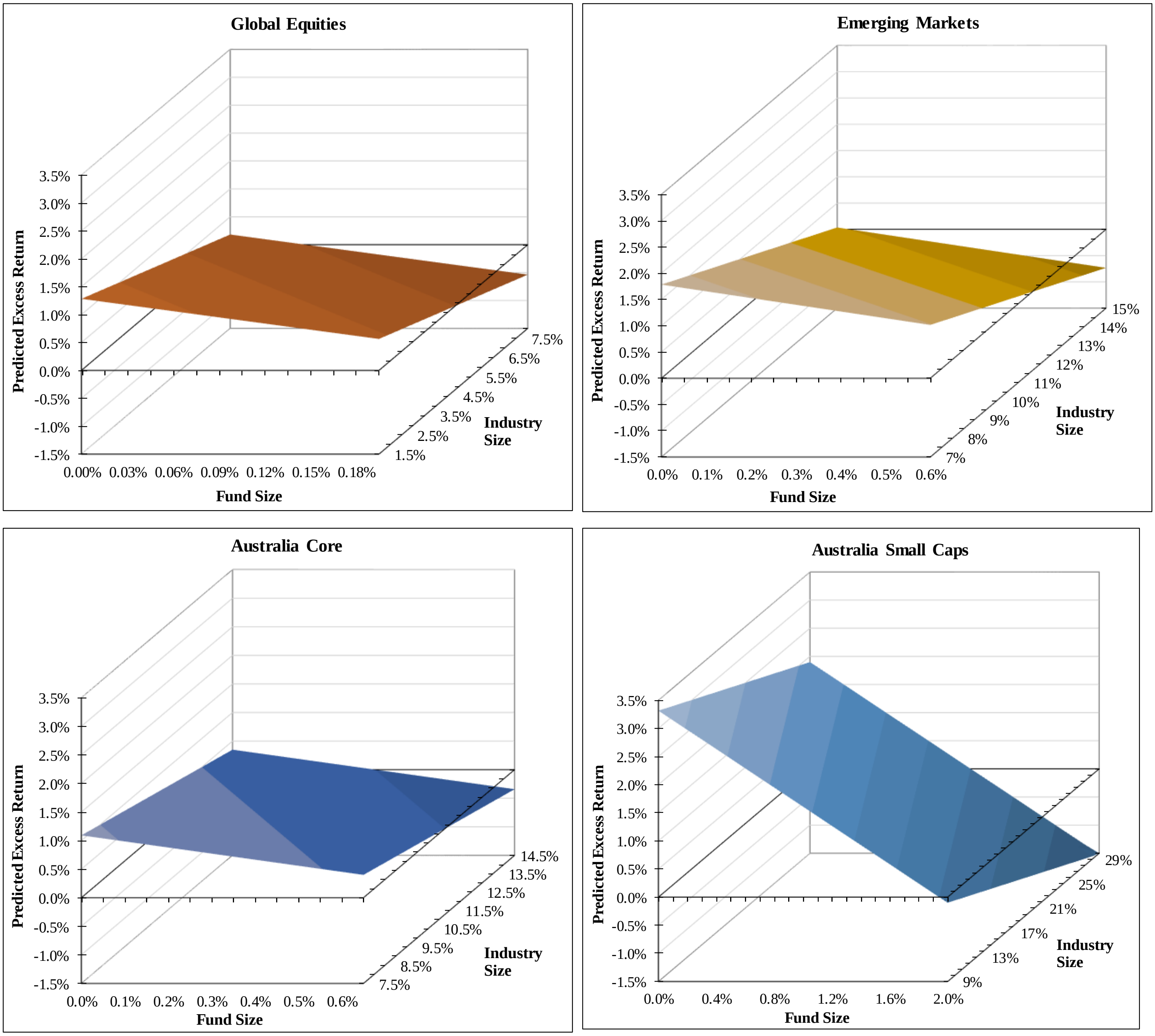
	Global Equities	Emerging Markets	Australia Core	Australia Small Caps
PANEL A: Linear Model				
Fund Size (FS)	-3.69	-1.28	-1.06	-1.71
t-statistic	-4.38 ***	-3.32 ***	-3.59 ***	-5.45 ***
Industry Size (IS)	-0.17	-0.22	-0.11	-0.07
t-statistic	-10.75 ***	-14.46 ***	-3.65 ***	-3.58 ***
R-squared	0.8%	2.4%	0.8%	2.2%
No. of Observations	19,147	9,733	3,865	2,365
PANEL B: Quadratic Model				
Fund Size (FS)	-8.96	-2.98	-2.29	-3.96
t-statistic	-5.26 ***	-4.55 ***	-4.37 ***	-6.11 ***
Fund Size Squared (FS^2)	1,263.5	117.0	68.2	102.7
t-statistic	3.41 ***	3.28 ***	2.79 ***	3.88 ***
Industry Size (IS)	0.31	-0.80	0.21	0.36
t-statistic	3.24 ***	-4.34 ***	0.62	2.85 ***
Industry Size Squared (IS^2)	-4.46	2.62	-1.43	-1.09
t-statistic	-5.04 ***	3.19 ***	-0.91	-3.37 ***
R-squared	1.0%	2.6%	1.0%	3.3%
No. of Observations	19,147	9,733	3,865	2,365
PANEL C: Data Range				
Fund Size				
1st Percentile	0.000042%	0.000334%	0.00028%	0.0030%
99th Percentile	0.219%	0.572%	0.708%	1.876%
Industry Size				
1st Percentile	2.06%	7.54%	7.92%	9.38%
99th Percentile	7.88%	14.68%	14.22%	29.36%

***/**/* indicates significance at the 1%/5%/10% level

To help interpret the regression estimates, surface plots are generated of the predicted relation between *XR* and both *IS* and *FS* over the observed data ranges, imposing a minimum *FS* of zero. The intercept is calibrated so that predicted *XR* equals the average *XR* when both *FS* and *IS* are also at their average for the market. The plots thus show how predicted *XR* varies with *FS* and *IS* for an average fund with a baseline *XR* in line with the market average. The vertical *XR* axes are scaled to span consistent ranges to assist visual comparability across markets. Figure 1 plots the linear models, with quadratic models provided in the Appendix as Figure A1. The plots indicate the surface for a 'typical' fund, bearing in mind that potential to generate returns (i.e. the intercept) varies across funds and fees need to be taken into account to arrive at *NXR*. The time series of *NXR* is examined in Section 4.4.

Figure 1: Relation Between AUM, Industry Size and Fund Size – Linear Models

Figure 1 plots the surface of predicted excess return (*XR*) under the linear regression models for each market. The intercept is calibrated so that predicted *XR* equals the average *XR* for the sample in each market when both fund size (*FS*) and industry size (IS) are also at their average. Each plot reflects the observed data range for *FS* and *IS* in each market, while imposing a minimum *FS* of zero. The value where the curve touches the y-axis can be interpreted as the predicted *XR* for *FS* of zero and an *IS* equal to the minimum observed in each market sample.



The linear model plots appearing in Figure 1 indicate that *XR* declines with both *FS* and *IS*, and that gross *XR* tends to become negative at the upper end of the *FS* and *IS* range within each market. One notable feature is that the *XR* surface for Australia small caps starts with a relatively high *XR* but is more steeply sloped as *FS* and *IS* increase. This indicates that underlying potential to generate *XR* is greater for this market, but so too is the impact of increasing AUM on the ability to sustain returns, i.e. diseconomies of scale are quite strong. This finding aligns with Australia small caps being the smallest of the markets being examined. As at the end of March 2017, the MSCI All Country market capitalization was approximately US$40.0 trillion, MSCI Emerging Markets US$4.4 trillion, S&P/ASX300 $US1.2 trillion, and Australia Small All Ordinaries only $US117 billion.

With regard to the quadratic models, the plots appearing in Figure A1 in the Appendix indicate that *XR* broadly declines with *FS* across the observed range for all markets, with the slope of the surfaces being not dissimilar to the linear models. The motivation behind including quadratic models was to gauge if economies of scale may be occurring before diseconomies of scale kick-in at higher AUM. The regression estimates provide no indication that this is occurring. Meanwhile, the quadratic models demonstrate a number of features that call their reliability into question. First, inconsistencies emerge in the non-linearities across markets that are hard to explain. The relation between *XR* and *IS* demonstrates a shallow hump (i.e. concave relation) for global equities and Australia core, a marked hump for Australia small caps, and a saucer shape (i.e. convex relation) for emerging markets. Meanwhile, all coefficients on the squared term for *FS* are positive, suggesting that continued increases in *FS* will eventually lead to higher *XR* beyond the existing sample range. We can offer no strong economic explanations for these patterns. Consideration needs to be given to the possibility that the quadratic estimates could be somewhat spurious, given potential for quadratic terms to be relatively more sensitive to the data. After reviewing the fit of models[5] and considering the plausibility of the predicted values, we decided to use the more parsimonious linear models for the remainder of the analysis.

The regression results are consistent with scale diseconomies across all markets examined, along with a role being played by both the size of the fund itself and the total amount of AUM competing in the market. This is consistent with the findings of Zhu (2018) for US equities. It is worth contemplating what might be the drivers behind these results, and the variation observed across markets. While a variety of potential sources of capacity constraints exist at the fund level (see O'Neill and Warren, 2019), two are particularly notable. First is that funds may face decreasing returns as AUM increases due to greater implementation shortfall (see Perold, 1988), related to a deterioration in the trade-off between execution costs (i.e. market impact) and opportunity costs (i.e. failing to trade) as larger trades become required to establish positions. Second, as AUM increases, the ability to establish positions in smaller stocks can be constrained by limits on the percentage of the market capitalization that a fund may hold (see O'Neill, Schmidt and Warren, 2018; O'Neill and Warren, 2019). Both implementation shortfall and holding constraints are more likely to bind in markets containing smaller and less liquid stocks. Returns can be further eroded by growth in industry AUM due to increasing competition for opportunities (Pástor and Stambaugh, 2012; Pástor, Stambaugh and Taylor, 2015). These explanations suggest that we should observe greater sensitivity to higher AUM in smaller markets. The estimates for Australia small caps are strongly consistent with these contentions. However, the lack of meaningful differentiation across the other three markets does not provide clear support for this contention. For instance, global equities might be expected to show less signs of scale diseconomies than emerging markets and Australia core due to the size and scope of global equity markets. We conclude that we provide strong evidence of scale diseconomies across multiple markets related to both fund size and industry size, but only tentative evidence that capacity constraints might bind more quickly in smaller markets with reference to the results for Australia small caps.

The regression models and Figure 1 speak to the 'average' fund. Table 3 reveals the distribution *XR* across funds through extracting values at selected percentiles. Statistics for realized *XR* are reported in Panel A, while predicted *XR* values under the linear models are reported in Panel B. Predicted *XR* spans a much narrower range than realized *XR*, suggesting that the model is doing an effective job at explaining some of the variability but is some way from explaining more extreme values in realized *XR*. Nevertheless, there remains meaningful variation in the predicted values, with the difference between the 90th and 10th percentiles ranges span from 0.9% for Australia core to 2.5% for Australia small caps. Table 3 suggests that realized *XR* may deviate considerably from predicted *XR*, hinting that some funds could be operating away from their predicted capacity. This matter is explored in Section 4.4.

Table 3: Distribution of Gross Excess Returns – Realized and Predicted

Table 3 uses selected percentiles to convey the distribution of realized excess return (*XR*) in Panel A and predicted *XR* in Panel B. Predicted *XR* is formed by combining the slopes estimated from the linear regression as reported in Panel A of Table 2 with observation of fund size (*FS*) and industry size (*IS*) in each quarter, and adjusting for fund-specific intercepts.

[5] Akaike information criterion (AIC) and Bayesian information criterion (BIC) scores revealed no meaningful improvement to indicate any clear superiority of the quadratic models.

Percentile	1%	10%	25%	50%	75%	90%	99%	Interquartile Range	90th - 10th Percentile
PANEL A: Realized									
Global Equities	-6.7%	-2.5%	-0.9%	0.4%	1.8%	3.6%	9.1%	2.7%	6.1%
Emerging Markets	-6.1%	-2.2%	-0.7%	0.7%	2.3%	4.1%	10.4%	3.0%	6.3%
Australia Core	-5.5%	-1.7%	-0.5%	0.5%	1.5%	3.2%	7.9%	2.1%	4.9%
Australia Small Caps	-8.6%	-3.0%	-0.8%	1.8%	4.4%	7.5%	16.0%	5.2%	10.5%
PANEL B: Predicted									
Global Equities	-1.4%	-0.2%	0.2%	0.5%	0.8%	1.2%	2.2%	0.7%	1.4%
Emerging Markets	-0.8%	0.0%	0.4%	0.9%	1.4%	1.8%	2.7%	1.0%	1.8%
Australia Core	-0.7%	0.2%	0.4%	0.6%	0.8%	1.0%	2.0%	0.4%	0.9%
Australia Small Caps	-0.8%	0.8%	1.3%	2.0%	2.7%	3.2%	5.4%	1.3%	2.5%

4.3. Ability of Models to Predict Excess Returns

To evaluate the regression models, we conducted robust regressions of realized *XR* against predicted *XR* for an out-of-sample period comprising one-third of the entire sample period[6]. Predicted returns are formed by estimating the models using an expanding data window starting at the two-thirds point in the sample (i.e. after the quarter 38), which is re-estimated each quarter until the end of the sample period (i.e. quarter 57). Predicted *XR* for each fund in the forthcoming quarter is formed by conditioning on observed *FS* and *IS* at the end of the prior quarter, applying the estimated slope coefficients and fund-specific intercepts at that point. The predicted *XR* for each quarter are regressed on the predicted *XR* as estimated at the end of the prior quarter.

Results are reported in Table 4. An indication that our models are working perfectly in predicting *XR* conditional on *FS* and *IS* would be an intercept of zero and a slope coefficient of one. Slope estimates are positive and significant with the notable exception of emerging markets, but are substantially less than one. Intercepts are significantly greater than zero. These results are consistent with the linear model explaining some portion of the observed *XR* during the last third of the sample for all but emerging markets, i.e. they contain some predictive information. Nevertheless, low predictive power of out-of-sample future fund returns is to be expected given the noisy nature of equity markets. The unexplained variation might also be partly due to presence of omitted variables not captured by the fund-specific intercepts.

[6] As ordinary least squares estimates are well known to be sensitive to outliers, we undertook robust regression analysis that reduces the influence of outliers. Several recent fund studies advocate use of robust regressions (e.g. Pástor et al., 2022; Adams et al., 2018). We use the M-estimator of Huber (1964), which minimizes a loss function that is quadratic for small residuals but linear for large residuals.

Table 4: Regression of Realized on Out-of-Sample Predicted Excess Returns

Table 4 reports results from regressing realized excess return in quarter *t* on predicted *XR* for each fund estimated by conditioning on fund size (*FS*) and industry size (*IS*) observed as at quarter *t-1*. Predictions of excess return are based on expanding data windows starting from quarter 38 (i.e. two-thirds through sample) through to quarter 57. This test indicates the extent to which the model contains information that assists in predicting fund *XR* in the latter part of the sample period. An intercept of zero and a slope coefficient of one represents perfect predictive ability.

	Global Equities	Emerging Markets	Australia Core	Australia Small Caps
Intercept	0.066	0.475	0.268	0.579
t-stastistic	2.43 **	11.78 ***	3.07 ***	2.19 **
Slope	0.219	0.041	0.311	0.392
t-statistic	6.59 ***	0.75	2.92 ***	4.43 ***
1 - Slope	0.781	0.959	0.689	0.608
F-test #	39.98 ***	0.54	8.38 ***	18.89 ***
df	1, 8307	1, 4155	1, 1496	1, 918

F-test, i.e., a Wald test for multiple coefficients

Predictions of excess return are based on the models estimated using expanding data windows from quarter 38 (i.e. two-thirds through sample). Predictions condition on oberved *FS* and *IS* at end of prior quarter.

4.4. Fund Scale versus Capacity

We now investigate the extent to which funds operate relative to their predicted capacity, inspired by Zhu (2018) and Barras, Gagliardini and Scaillet (2022) as well as Roussanov, Ruan and Yanhao (2021). The analysis is facilitated by applying equation (6) to estimate predicted net excess return (i.e. $\widehat{NXR}$) for each fund in each quarter, conditional on observed *FS*, *IS* and a 'breakeven' hurdle rate of return h_i that equates with fund fees. Funds with a negative $\widehat{NXR}$ are interpreted as operating at an AUM consistent with being above capacity, and vice versa. Initially we estimate the percentage of funds operating at above predicted capacity in each market for each quarter, and relate trends in these series to *FS* and *IS*. We subsequently apply equation (10) to establish confidence intervals and thus identify funds that are significantly above or below their predicted capacity. These results are used to examine the pervasiveness and persistence of significant deviations from the BG equilibrium. Applying a hurdle rate of return (h_i) reflecting fund fees also effectively applies the terminal capacity (i.e. zero net alpha) definition of capacity as proposed by Vangelisti (2006).

There is no one level for fund fees, which may vary across markets, distribution channels (retail, wholesale and institutional mandates), mandate size and individual funds (possibly reflecting skill or bargaining power). We thus use the Mercer fee report (Mercer, 2020) to form an indicative 'low' and 'high' fee for each market for the purpose of these tests, with a view to generating a representative range for $\widehat{NXR}$. The low fee reflects quartile 1 fees for segregated mandates, which captures the lesser fees paid by larger institutional investors. The high fee reflects quartile 3 fees for retail mutual funds, which typically pay greater fees. Fee data is extracted with reference to the fund category reported by Mercer that is deemed as most representative of the market in question. Table 5 present the fee assumptions we apply in each market.

Table 5: Fee Assumptions for Capacity Estimation

Table 5 presents low and high fee assumptions for each market. The assumed fees are based on data from Mercer (2020). The low fee reflects quartile 1 fees for segregated mandates and the high fee reflects quartile 3 fees for retail mutual funds, using the fund category reported by Mercer deemed as most representative.

	Low Fee	**High Fee**	
Investment of US$25 million in:	*Segregated Mandate*	*Retail Mutual Fund*	***High - Low***
Fee struck at:	*Quartile 1*	*Quartile 3*	***Difference***
Global Equities	0.60%	1.21%	0.61%
Emerging Markets	0.75%	1.30%	0.55%
Australia Core	0.41%	1.33%	0.92%
Australia Small Caps #	0.83%	1.54%	0.71%

As fee data was only available for wholesale small cap funds, we adjusted this data for the median differences between segregated mandates and retail mutual funds versus wholesale funds as reported for Australia Core.

4.4.1. Analysis of Percentage of Funds Above Capacity

Figure 2 plots time series of the percentage of total AUM where $\widehat{NXR_{\iota,t}}$ is negative at both a high fee and a low fee. These series indicate the percentage of fund AUM that is estimated to be operating above capacity in each quarter, i.e. 'over-funded', in the sense of managing too much AUM. Time series plots of the percentage of funds by number operating above capacity are provided in the Appendix as Figure A2. As the plots reflect a hard cut-off of $\widehat{NXR_{\iota,t}} < 0$, the complement may be interpreted as funds that are operating at or below capacity. Significant deviations from $\widehat{NXR_{\iota,t}} = 0$ and hence capacity are investigated in Section 4.4.2.

For both global equities and emerging markets, the percentage of AUM operating above capacity starts near zero and then increases through the sample period to around the 10%-50% range. A jump occurred for global equities in about 2007-2009 and for emerging markets around 2013-2014. For Australia core and Australia small caps, the percentage exceeding capacity by both AUM and number of funds starts at a high level and then declines, before rising again towards the end of the period. The percentage of AUM above capacity is moderately greater at a high fee than at the low fee by around 15% or less for the most part, suggesting that the sensitivity of capacity estimates to the fee assumption is mostly moderate. However, sensitivity to the assumed fee can be substantial on occasion. For example, larger differences in the percentage above capacity between low and high fees are more commonplace for Australia core. The trends in funds above capacity by number (Appendix, Figure A2) are broadly similar, with the exception of global equities where this series trends progressively higher after 2009 while the 'percentage of AUM' series stabilizes. Indeed, around 35%-50% of the global equities fund sample by number is estimated to be above capacity near the end of the sample period, versus 20%-35% by AUM. The key takeaway is that movements in the percentage of fund operating above capacity vary noticeably both through the sample period and across markets.

Figure 2: Percentage of AUM Exceeding Capacity (i.e. $\widehat{NXR} < 0$)

Figure 2 plots the time series of the estimated percentage of fund AUM that is operating above capacity at a high and low fee in each market, as indicated by a negative value for predicted net excess return (i.e. $\widehat{NXR}_{i,t}$). The latter is estimated using equation (6), and involves combining the slopes estimated from the linear regression as reported in Panel A of Table 2 with observations of fund size (*FS*) and industry size (*IS*) in each quarter, and then adjusting for fund-specific intercepts and the assumed fee.

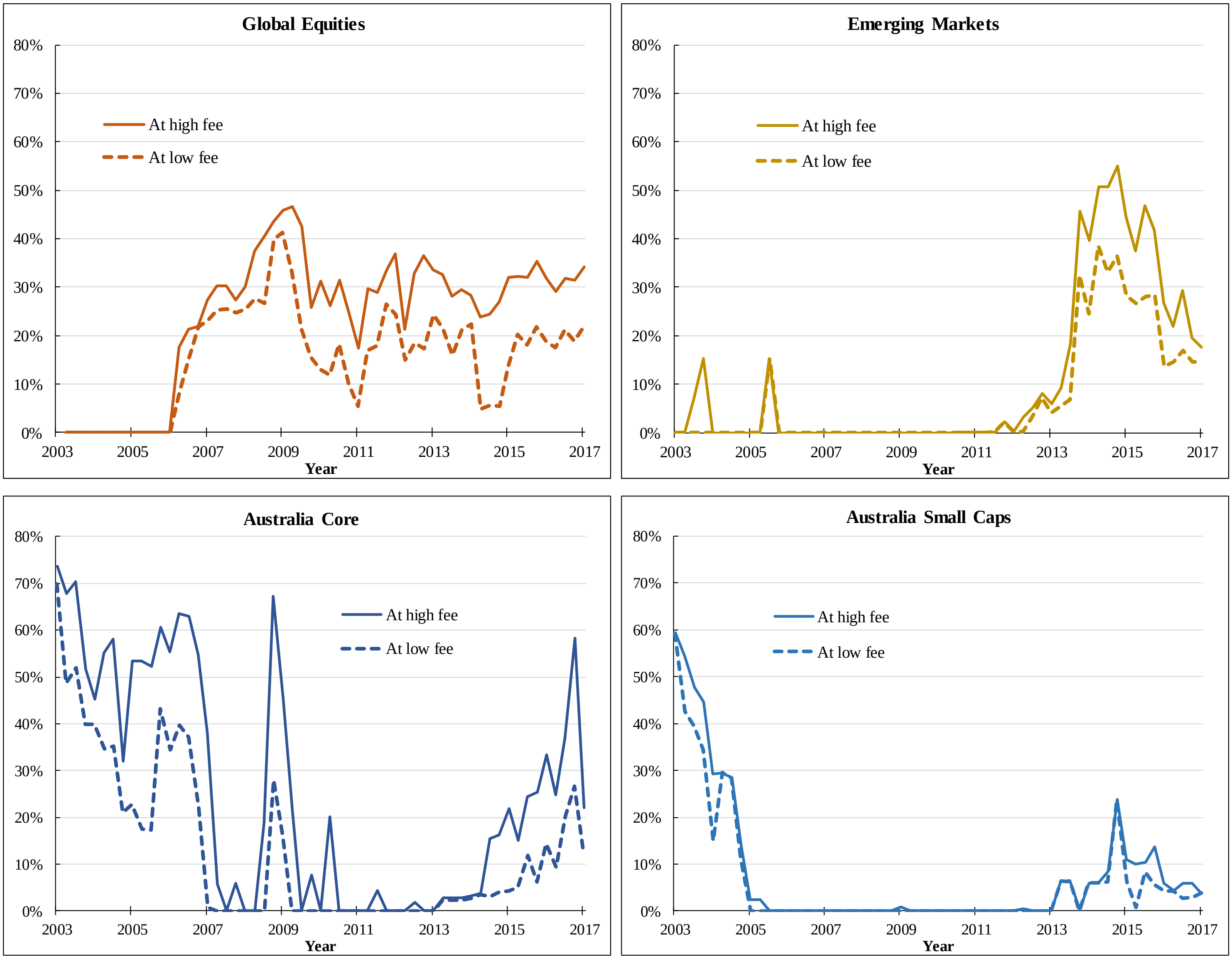


Figure 3 investigates the drivers behind Figure 2 by plotting the percentage of AUM exceeding capacity at a high fee against *IS* (charts on the left) and mean *FS* (charts on the right). These charts reveal some notable differences in the role of *FS* and *IS* across markets and across time. For global equities, both *IS* and mean *FS* rise through to 2007-2009, suggesting that growth in industry AUM tended to be directed towards increasingly larger funds. After that, *IS* continues to increase while mean *FS* decreases, suggesting a new phase where additional industry AUM was being spread among more funds. It is possible that this could have occurred in recognition that some larger global equity funds were hitting their capacity constraints. Emerging markets reveal a progressive increase in *IS* and decrease in mean *FS* throughout the sample period. This pattern is consistent with the capacity constraints that appear towards the end of the sample period being primarily driven by an increase in industry AUM, with this AUM being spread across more funds. For Australia core, relatively high *IS* and mean *FS* appear to explain the relatively high level of AUM above capacity at the start of the sample, with the fluctuations thereafter aligning with trends in *IS* while mean *FS* declines. This is consistent with *IS* being the primary driver, while available AUM is being split between more funds over time. For Australia small caps, both *IS* and mean *FS* appear to be correlated with changes in the percentage of AUM operating above capacity at various times, suggesting that both factors may have been playing a role. The sharp decline in AUM operating above capacity early in the sample period is associated with a sharp decline in mean *FS*, suggesting that capacity constraints in Australia small caps were being addressed through reallocation of AUM across funds during the early period. The increase in AUM above capacity towards the end of the sample period is associated with rises in both *IS* and mean *FS*.

Figure 3: Percentage of AUM Exceeding Capacity versus Industry Size and Fund Size

Figure 3 plots the time series for each market of the estimated percentage of fund AUM that is operating above capacity at a high fee (as indicated by a negative value for predicted net excess return, i.e. $\widehat{NXR}_{i,t}$) against industry size (charts on left side) and average fund size (charts on right side).

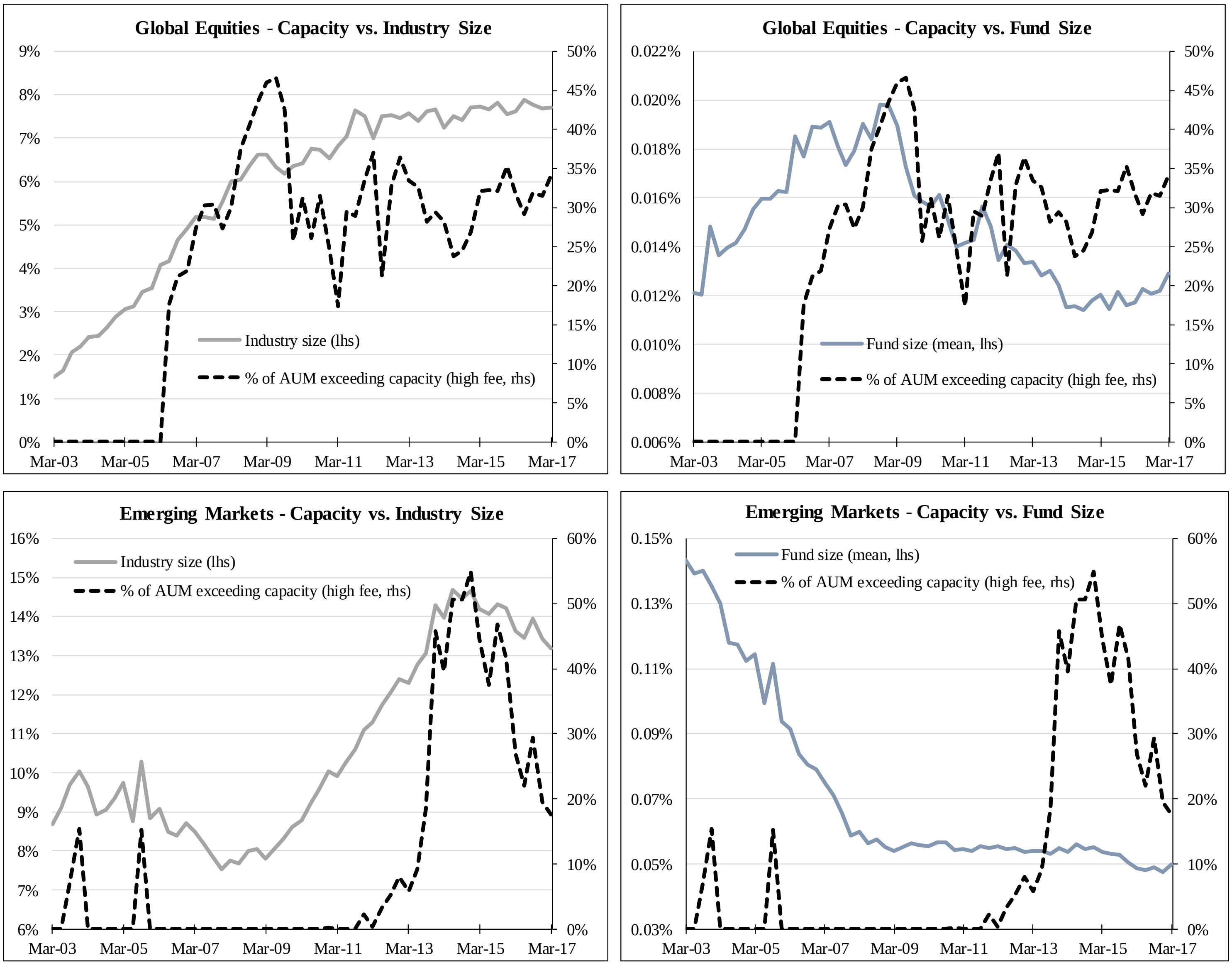
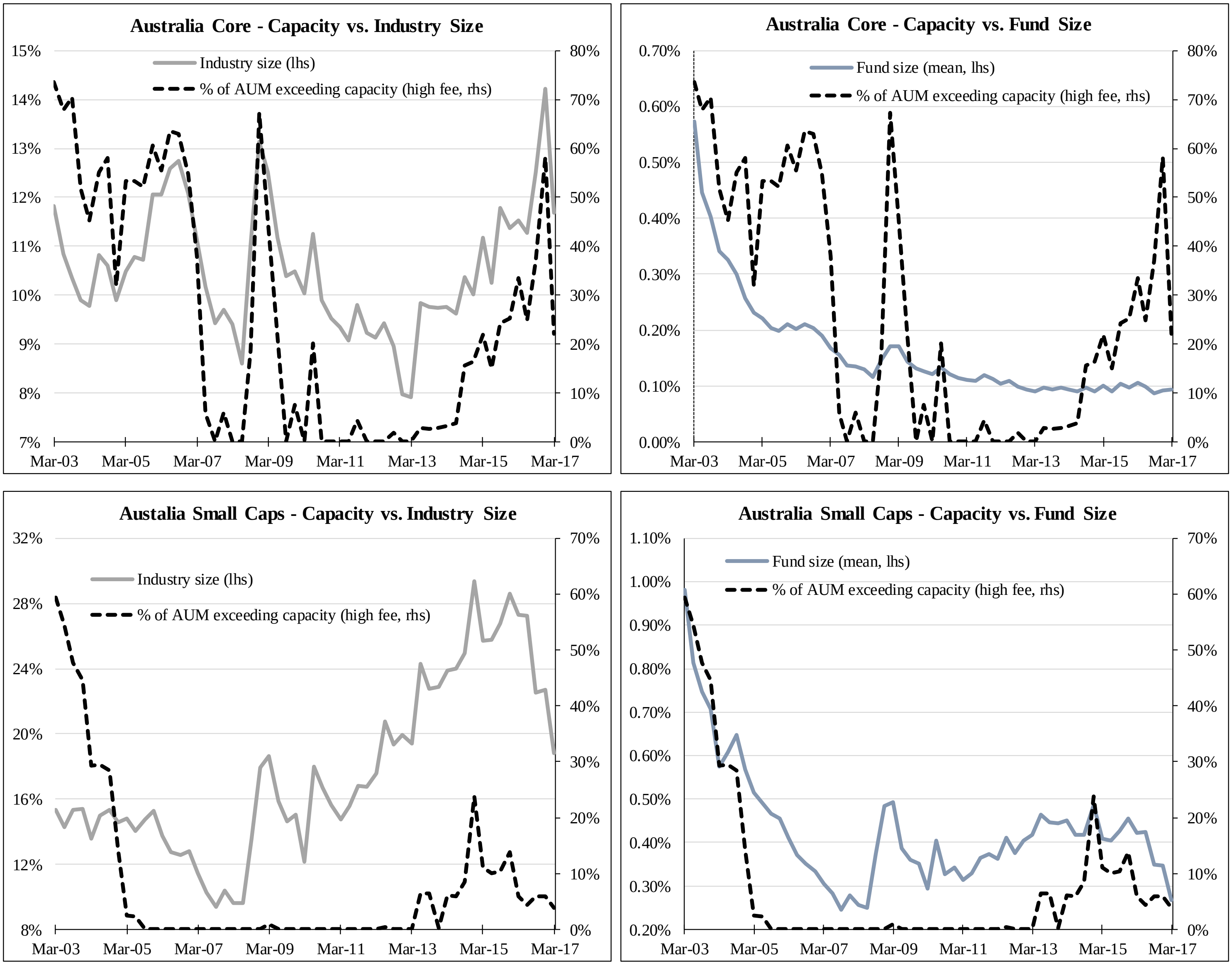


Figure 3: Percentage of AUM Exceeding Capacity versus Industry Size and Fund Size *(continued)*



The analysis of the percentage of AUM and number of funds operating above capacity, along with the examination of the relation with *FS* and *IS*, suggests that market structure and dynamics are playing important roles in determining the extent to which capacity constraints are encountered within a market. Capacity constraints may become more prevalent due to either increases in AUM at the industry level as more assets are directed towards active funds, how that AUM is allocated across funds, or a combination of both. Further, the impact of these influences appears to vary across both markets and across time. The results support two conclusions. First, it is important to consider both *IS* and *FS* in explaining capacity movements, which is consistent with Zhu (2018). Second, the dynamics behind changes in the portion of funds operating above versus below capacity can differ substantially across markets.

4.4.2. Analysis of Funds Significantly Above and Below Capacity

The previous sub-section examined the percentage of AUM operating above capacity, as identified by a negative $\widehat{NXR_{i,t}}$ value. While this analysis reveals broad trends within each market, clear evidence of inconsistency with the equilibrium described by BG requires taking into account statistical significance. We do so by calculating confidence intervals through estimating the variance of $\widehat{NXR_{i,t}}$ using equation (10). Our tests are based on 80% confidence intervals (i.e. 10% and 90% p-values), as a compromise between generating sufficient observations for a meaningful analysis while still applying a respectable level of statistical significance. The standard errors are relatively large as a consequence of both small sample sizes for some funds, and the possibility that our models may not capture all factors that may be relevant for fund-level analysis.[7] We also estimated results at the more standard 90% confidence interval, but this produced too few observations of funds with significantly positive or negative $\widehat{NXR}$ to support a useful analysis of trends and persistence of deviations from the BG equilibrium. We focus on the percentage of funds by number for this analysis, which is directed at gauging how effectively the market is able to identify capacity at the individual fund level and adjust accordingly. Figure 4 reports the results as time series plots.

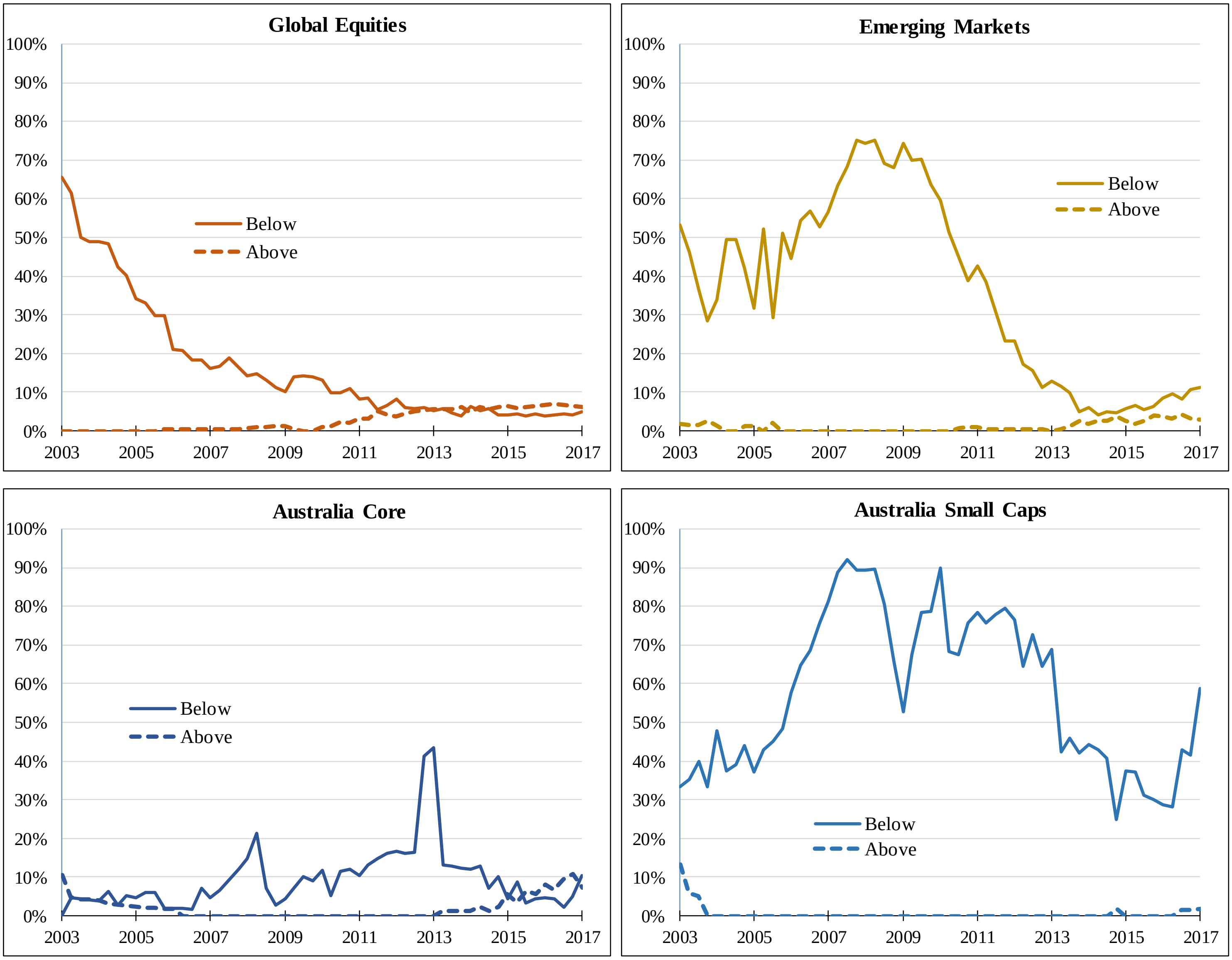
One notable finding from Figure 4 is that far more funds are estimated to be operating significantly below capacity than significantly above capacity. Indeed, the percentage of funds estimated to be operating significantly above capacity is typically less than 10%, which would be the portion expected by random variation alone. In contrast, the percentage estimated to be operating significantly below capacity is often well in excess of 10%. This is most notable for global equities and emerging markets until the early-2010's, and for Australia small caps throughout the sample period. The results are broadly consistent with the BG equilibrium for Australia core, and for global equities and emerging markets in the latter few years of the sample period. Conducting the analysis assuming low fees rather than high fees acts to exacerbate the skew towards more funds being significantly below capacity, as it lowers the hurdle (h) applied to XR in estimating $\widehat{NXR}_{i,t}$.

There are a number of potential explanations for the observed propensity for more funds to be operating significantly below capacity in the markets we analyze. One possibility is frictions in identifying skilled managers and adjusting flows and hence AUM, so that $\widehat{NXR}$ is eroded only at a lag. This could occur due to learning effects, where investors observe positive XR over a period of time before gaining confidence that skill exists and flows respond, e.g. see Foster and Warren (2015); Yan, 2019; Barras, Gagliardini and Scaillet (2022). Another possibility is that IS may be the primary driver at times, and may reflect influences other than attraction to manager-specific skill, e.g. asset allocation decisions. The trends observed for global equities and emerging markets, where the percentage of funds significantly below capacity decreases and IS rises over the sample period, could be consistent with such learning and industry-wide influences. However, these explanations are harder to square with the results of Australia small caps, as well as Australia core. Another possibility is estimation issues. For example, 'true' capacity may be either over-estimated by our model or under-estimated by investors, either of which would result in less AUM being allocated than the level at which $\widehat{NXR} = 0$ under our model. Another possibility is that investors have some bargaining power and require $\widehat{NXR} > 0$ to allocate funds, i.e. their hurdle (h) exceeds the fee. Finally, there could be behavioral influences, such as the propensity for some skilled funds to be overlooked by investors due to inattention or poor marketing efforts.

[7] Capacity constraints at the fund level might reflect fund-level attributes other than AUM that we are unable to observe. For instance, Pástor, Stambaugh and Taylor (2020) argue that scale depends on not only AUM but also activeness, which is a function of turnover and liquidity. O'Neill, Schmidt and Warren (2018) highlight how the relation between AUM and capacity constraints may differ with investment process.

Figure 4: Percentage of Funds by Number Significantly Below and Above Capacity at High Fee

Figure 4 plots the time series for each market of the estimated percentage of funds by number that are operating significantly above or significantly below capacity, as indicated by a negative value for predicted net excess return (i.e. $\widehat{NXR}_{i,t}$) assuming a high fee. The charts reflect 80% confidence intervals (e.g. percentage of funds below the 10% and above the 90% p-values), which is estimated using equation (10).



In any event, a notable finding is that the portion of funds that are significantly below rather than above capacity differs across markets, suggesting that market-specific drivers are at play. Not only do our results vary across the four markets examined, but they also contrast with those for US equity mutual funds where studies find relatively more funds operating significantly above than significantly below capacity, e.g. Zhu (2018); Song, (2020); Roussanov, Ruan and Yanhao Wei (2021); Barras, Gagliardini and Scaillet (2022).

We extend the analysis of significant deviations from capacity by investigating its persistence. Table 6 reports the mean and selected percentiles for the number of quarters that individual funds remain significantly below capacity (Panel A) and significantly above capacity (Panel B), which we refer to as an 'episode'. Some caution needs to be applied in interpreting the estimates in some instances due to low sample size. In particular, we only observe four episodes where an Australia small cap fund spent time operating significantly above capacity, and only 17 such episodes for Australia core. These results provide an indication of the speed at which significant deviations from the BG equilibrium are addressed, as revealed by $\widehat{NXR}_{i,t}$ moving back within the 80% confidence interval. As the distribution of quarters significantly above or below capacity is right-skewed, we will first focus on the medians rather than the means before discussing the right tail of estimates.

Table 6: Distribution of Quarters Spent Significantly Below and Above Capacity (High Fee)

Table 6 reports the mean and selected percentiles for the number of quarters that particular funds remain significantly below (Panel A) and significantly above (Panel B) capacity, which we refer to as an episode. Above capacity funds are identified by a significantly negative value for predicted net excess return (i.e. $\widehat{NXR}_{i,t}$) assuming a high fee at a 10% confidence level, with the latter estimated using equation (10). Similarly, below capacity funds are identified by a significantly positive value for $\widehat{NXR}_{i,t}$. The length of an episode is measured by identifying the quarter when a fund is observed to move either significantly above- or below-capacity, and then counting the number of quarters until $\widehat{NXR}_{i,t}$ subsequently becomes statistically insignificant. This data is used to estimate the mean and as percentiles reported.

Significance cut-offs: 10% and 90%	**Sample Size**		**Mean**	**Percentiles**				
	Fund/Qtrs	**Episodes**		**10%**	**25%**	**50%**	**75%**	**90%**
Panel A: Quarters Below Capacity								
Global Equities	888	225	9.5	1	2	4	12	24
Emerging Markets	689	267	10.5	1	2	6	18	27
Australia Core	245	74	5.2	1	1	2	4	16
Australia Small Caps	244	127	10.6	1	1	5	16	30
Panel B: Quarters Above Capacity								
Global Equities	699	75	9.5	1	2	6	15	25
Emerging Markets	364	30	4.6	1	1	3	7	9
Australia Core	144	17	4.8	1	1	2	7	12
Australia Small Caps	82	4	2.0	1	1	2	3	3

The median time that funds spend either significantly below or significantly above our capacity estimates range from two to six quarters across the markets. This suggests that significant deviations from the BG equilibrium often do not persist for an overly extended period of time. A period of adjustment of two-six quarters does not seem unreasonable given the difficulty in observing manager skill. However, the 90th percentiles indicate that substantial deviations can persist for two years or more for a minority of funds, most notably in global equities for fund both below- and above-capacity, and below-capacity funds in emerging markets and Australia small caps. This might reflect extreme cases of the potential influences raised in discussing Figure 4. However, it is perhaps more likely that these instances could stem from capacity being poorly estimated for some funds. For instance, consider a situation where our model over-estimates capacity by failing to fully account for (say) the activeness of the strategy (see Pástor, Stambaugh and Taylor, 2020), which is recognized by investors who limit AUM to below the capacity implied by our model. The result could be a situation where our model persistently estimates capacity to be well above the level of AUM the investors are (correctly) willing to supply. To the extent that this is occurring, it is comforting that it might apply only to a modest portion of our sample.

Finally, the results reported in Table 6 provide further support for the theme of differences across markets, in this case with respect to dynamics of AUM adjustment. The extent of these differences are summarized by the variation in the mean number of quarters that funds spend significantly below or above capacity. Putting aside the below-capacity estimates for Australia small caps where there are only 4 episodes, the means range between 4.6 and 10.6 quarters, i.e. a range of 1.5 years.

5. Conclusion

We have related *XR* to *FS* and *IS* in four equity markets, and used the resulting models to identify and analyze deviations from capacity with reference to whether predicted *NXR* (i.e. excess returns after fees) are positive or negative. The findings support two contentions. First is that diseconomies of scale are evident across a selection of equity markets, and are related to both *FS* and *IS*. Our findings add to similar evidence for US equity mutual funds, most notably Zhu (2018). Second is that the extent to which funds operate near capacity and the dynamics of adjustment appear to differ across markets. We uncover notable differences across the four equity markets examined in the prevalence of funds that are operating below- or above-capacity, the speed of adjustment towards capacity, and the relative influence of *FS* and *IS* in driving market-wide trends. We also find a propensity for substantially more funds to be significantly below-capacity in the markets examined, which contrasts with US results that point towards more funds operating above-capacity. This implies that there may be more scope to access positive excess returns from active management outside of the US, which aligns with findings from the literature such as Gallagher et al. (2017) and Lieppold and Rueegg (2020) for global equities, and Chen et al. (2010) and Cao, von Reibnitz and Warren (2020) for Australia. The findings that capacity dynamics differ across markets cautions against extrapolating results for US equity funds to other markets.

We consider our findings as indicative rather than conclusive. There are numerous econometric hurdles in estimating economies of scale and capacity in fund management, including untangling the endogeneity between AUM and fund returns while accounting for all potential drivers of scale diseconomies. For instance, we rely on fund fixed effects to capture relevant variations in potential to generate excess returns, while assuming that there is a common relation between variation in AUM and returns across funds. To the extent that these assumptions are not correct, our fund-level results may be subject to impacts from omitted variables as well as noise. The fund data from Mercer's GIMD™ database could contain some selection bias; and our *IS* proxy is rough. Finally, we examine only four 'markets' spanning two international and two Australian equity categories. Expanding the tests to other markets and perhaps other assets presents a future research direction. Nevertheless, we believe that we are on relatively safe ground with regard to our main findings, which are twofold. First, scale diseconomies in fund management can be found in a range of equity markets that relate to both fund size and industry size. Second, the nature of the relation between AUM, capacity and excess returns can vary substantially across markets and over time.

References

Adams, John, Darren Hayunga, and Sarrar Mansi (2018). Diseconomies of scale in the actively-managed mutual fund industry: what do the outliers in the data tell us? *Critical Finance Review* 7(2): 373—377.

Ammann, Manuel, and Patrick Moerth (2008). Impact of fund size and fund flows on hedge fund performance. *Journal of Alternative Investments*, 11(1): 78-96.

Andonov, Aleksandar, Rob Bauer and Martijn Cremers (2012). Can large pension funds beat the market? Asset allocation, market timing, security selection and the limits of liquidity. *Netspar Discussion Paper*, No. 10/2012-062. (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2214931)

Andonov, Aleksandar, Piet Eichholtz and Nils Kok (2015). Intermediated investment management in private markets: Evidence from pension fund investments in real estate. *Journal of Financial Markets*, 22: 73-103.

Barras, Laurent, Patrick Gagliardini and Olivier Scaillet (2022). Skill, scale, and value creation in the mutual fund industry. *Journal of Finance*, 77(1): 601-638.

Berk, Jonathan B. and Richard C. Green (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy*, 112(6): 1269-1295.

Cao, Ying, Anna von Reibnitz, and Geoffrey J. Warren (2020). Return dispersion and fund performance: Australia – The land of opportunity? *Pacific-Basin Finance Journal*, 60: 101269.

Chan, Howard W., Robert W. Faff, David R. Gallagher and Adrian Looi (2009). Fund size, transaction costs and performance: Size matters! *Australian Journal of Management*, 34(1): 73-96.

Chen, Cong, Carole Comerton-Forde, David R. Gallagher and Terry S. Walter (2010). Investment manager skill in small-cap equities. *Australian Journal of Management*, 35(1): 23-49.

Chen, Joseph, Harrison Hong, Ming Huang and Jeffrey D. Kubik (2004). Does fund size erode mutual fund performance? The role of liquidity and organization. *American Economic Review*, 94(5): 1276-1302.

Dyck, Alexander and Lukasz Pomorski (2011). Is bigger better? Size and performance in pension plan management. *Working paper*, Rotman School of Management, University of Toronto. (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1690724).

Dyck, Alexander and Lukasz Pomorski (2016). Investor scale and performance in private equity investments. *Review of Finance*, 20(3): 1081-1106.

Elton, Edwin J., Gruber, Martin J. and Blake, Christopher R. (2012). Does mutual fund size matter? The relationship between size and performance. *Review of Asset Pricing Studies*, 2(1): 31-55.

Ferreira, Miguel A., Aneel Keswani, António F. Miguel, and Sofia B. Ramos (2013). The determinants of mutual fund performance: A cross-country study. *Review of Finance*, 17(2): 483-525.

Forsberg, David, David R. Gallagher and Geoffrey J. Warren (2022). Capacity constraints in hedge funds: The impact of cohort size on fund performance. *Financial Analysts Journal*, in press. (https://www.tandfonline.com/doi/full/10.1080/0015198X.2021.1996200).

Foster, F. Douglas and Geoffrey J. Warren (2015). Why might investors choose active management? *Journal of Behavioral Finance*, 16(1): 20-39.

Fung, William, Hsieh, David A., Naik, Narayan Y. and Ramadorai, Tarun (2008). Hedge funds: Performance, risk, and capital formation. *Journal of Finance*, 63(4): 1777-1803.

Gallagher, David R. and Kyle M. Martin (2005). Size and investment performance: A research note. *Abacus*, 41(1): 55-65.

Gallagher, David R., Graham Harman, Camille H. Schmidt and Geoffrey J. Warren (2017). Global equity fund performance: An attribution approach. *Financial Analysts Journal*, 73(1): 56-71.

Gao, Chao, Tim Haight and Chengdong Yin (2018). Size, age, and the performance life cycle of hedge funds. *Working paper*, Krannert School of Management, Purdue University. (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3169312).

Getmansky, Mila (2012). The life cycle of hedge funds: Fund flows, size, competition, and performance. *Quarterly Journal of Finance*, 2(1): 1-53.

Grinblatt, Mark and Titman, Sheridan (1989). Mutual fund performance: An analysis of quarterly portfolio holdings. *Journal of Business*, 62(3): 393-416.

Harvey, Campbell R., and Yan Liu (2017). Decreasing returns to scale, fund flows, and performance. *Duke I&E Research Paper*, No. 2017-13 (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2990737).

Hoberg, Gerard, Nitin Kumar, and Nagpurnanand Prabhala (2018). Mutual fund competition, managerial skill, and alpha persistence. *Review of Financial Studies*, 31(5): 1896–1929.

Huber, Peter J. (1964). Robust estimation of a location parameter. *Annals of Statistics* 53: 73—101.

Indro, Daniel C., Jiang, Christine X., Hu, Michael Y. and Lee, Wayne Y. (1999). Mutual fund performance: Does fund size matter? *Financial Analysts Journal*, 55(3): 74-87.

Kaplan, Steven N. and Antoinette Schoar (2005). Private equity performance: Returns, persistence, and capital flows. *Journal of Finance*, 60(4): 1791-1823.

Leippold, Markus and Roger Rueegg (2020). How rational and competitive is the market for mutual funds? *Review of Finance*, 24(3): 579-613.

Lopez-de-Silanes, Florencio, Ludovic Phalippou and Oliver Gottschalg (2015). Giants at the gate: Investment returns and diseconomies of scale in private equity. *Journal of Financial and Quantitative Analysis*, 50(3): 377-411.

McLemore, Ping (2019). Do mutual funds have decreasing returns to scale? Evidence from fund mergers. *Journal of Financial and Quantitative Analysis*, 54(4): 1683-1711.

Mercer (2020). *Global asset manager fee survey 2020*. Mercer, July.

Naik, Narayan Y, Tarun Ramadorai, and Maria Stromqvist (2007). Capacity constraints and hedge fund strategy returns. *European Financial Management*, 13(2): 239-256.

O'Neill, Michael J., Camille H. Schmidt, and Geoffrey J. Warren (2018). Capacity analysis for equity funds. *Journal of Portfolio Management*, 44(5): 36-49.

O'Neill, Michael J. and Geoffrey J. Warren (2019). Evaluating fund capacity: Issues and methods. *Accounting and Finance*, 59(S1): 773-800.

Pástor, Ľuboš. and Robert F. Stambaugh (2012). On the size of the active management industry. *Journal of Political Economy*, 120(4): 740-781.

Pástor, Ľuboš., Robert F. Stambaugh, and Lucian A. Taylor (2015). Scale and skill in active management. *Journal of Financial Economics*, 116(1), pp. 23-45.

Pástor, Ľuboš., Robert F. Stambaugh, and Lucian A. Taylor. (2020). Fund tradeoffs. *Journal of Financial Economics*, 138(3): 614-634.

Pástor, Ľuboš., Robert F. Stambaugh, Lucian A. Taylor, and Min Zhu, (2022). Diseconomies of scale in active management: robust evidence. *Critical Finance Review*, forthcoming.

Perold, André F. (1988) The implementation shortfall: Paper versus reality. *Journal of Portfolio Management*, 14(3): 4-9.

Phillips, Blake, Kuntara Pukthuanthong, and P. Raghavendra Rau (2018). Size doesn't matter: Diseconomies of scale in the mutual fund industry revisited. *Journal of Banking and Finance*, 88 (March): 357-365.

Philpot, James, Hearth, Douglas, Rimbey, James N. and Schulman, Craig T. (1998). Active management, fund size, and bond mutual fund returns. *Financial Review*, 33(2): 115-125.

Ramadorai, Tarun (2013). Capacity constraints, investor information, and hedge fund returns. *Journal of Financial Economics*, 107(2): 401-41.

Reuter, Jonathan and Eric Zitzewitz (2021). How much does size erode mutual fund performance? A regression discontinuity approach. *Review of Finance*, 25(5): 1395-1432.

Roussanov, Nikolai, Hongxun Ruan and Yanhao Wei (2021). Marketing mutual funds. *Review of Financial Studies*, 34(6): 3045-3094.

Song, Yang (2020). The mismatch between mutual fund scale and skill. *Journal of Finance*, 75(5): 2555-2589.

Stambaugh, Robert F. (2014). Presidential address: Investment noise and trends. *Journal of Finance*, 69(4): 1415-1453.

Vangelisti, Marco (2006). The capacity of an equity strategy. *Journal of Portfolio Management*, 32(2): 44-50.

Vidal-Garcia, Javier and Marta Vidal (2021). Short-term performance and mutual fund size. *Working paper*, Complutense University of Madrid. (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2801930).

Yan, Xuemin (2008). Liquidity, investment style, and the relation between fund size and fund performance. *Journal of Financial and Quantitative Analysis*, 43(3): 741-768.

Yan, Zen (2020). Returns to Scale Among Corporate Bond Mutual Funds, *Working paper*, Cornerstone Research. (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3339511).

Yin, Chengdong (2016). The optimal size of hedge funds: Conflict between investors and fund managers. *Journal of Finance*, 71(4): 1857-1894.

Zhu, Min (2018). Informative fund size, managerial skill, and investor rationality. *Journal of Financial Economics*, 130(1): 114-134.

APPENDIX

Table A1: Regression Estimates for Global Equities vs. Developed Markets and ACWI Funds

Table A1 presents results from regressing fund excess returns relative to the benchmark index (*XR*) against fund size (*FS*) and industry size (*IS*) while allowing for fund-specific fixed effects and hence intercepts, for ACWI and Developed Market funds. Regression results for global equities are reported for comparison. *IS* is estimated as the total sum of AUM for all the funds contained in the GIMD™ database that are addressing the market, divided by the total market capitalization for that market. *FS* is calculated by dividing the AUM for fund *i* by the market capitalization of its own investable market universe. Panel A reports results for regressions including linear terms only, while Panel B reports results for regressions include quadratic terms for both *FS* and *IS*. Panel C provides an indication of the data range by providing observations at the 1st and 99th percentiles for both *FS* and *IS*.

	Global Equities	**ACWI Funds**	**Developed Markets**
Benchmark	*MSCI World and All Country World Index*	*MSCI All Country World Index*	*MSCI World*
PANEL A: Linear Model			
Fund Size (FS)	-3.69	-3.27	-3.78
t-statistic	-4.38 ***	-1.96 *	-4.37 ***
Industry Size (IS)	-0.17	-0.38	-0.13
t-statistic	-10.75 ***	-5.12 ***	-6.14 ***
R-squared	0.8%	0.5%	0.6%
No. of Observations	19,147	7,663	11,410
PANEL B: Quadratic Model			
Fund Size (FS)	-8.96	-9.65	-8.57
t-statistic	-5.26 ***	-2.93 ***	-4.70 ***
Fund Size Squared (FS^2)	1,263.5	2,070.1	996.1
t-statistic	3.41 ***	2.38 **	2.97 ***
Industry Size (IS)	0.31	-1.27	-0.01
t-statistic	3.24 ***	-3.74 ***	-0.08
Industry Size Squared (IS^2)	-4.46	24.33	-1.27
t-statistic	-5.04 ***	2.78 ***	-0.85
R-squared	1.0%	0.7%	0.7%
No. of Observations	19,147	7,663	11,410
PANEL C: Data Range			
Fund Size			
1st Percentile	0.000042%	0.000050%	0.000043%
99th Percentile	0.219%	0.225%	0.226%
Industry Size			
1st Percentile	2.06%	0.62%	1.45%
99th Percentile	7.88%	3.09%	6.93%

***/**/* indicates significance at the 1%/5%/10% level

Figure A1: Relation Between AUM, Industry Size and Fund Size – Quadratic Models

Figure A1 plots the surface of predicted excess return (*XR*) under the quadratic regression models for each market. The intercept is calibrated so that predicted *XR* equals the average *XR* for the sample in each market when both fund size (*FS*) and industry size (IS) are also at their average. Each plot reflects the observed data range for *FS* and *IS* in each market, while imposing a minimum *FS* of zero. The value where the curve touches the y-axis can be interpreted as the predicted *XR* for *FS* of zero and an *IS* equal to the minimum observed in each market sample.

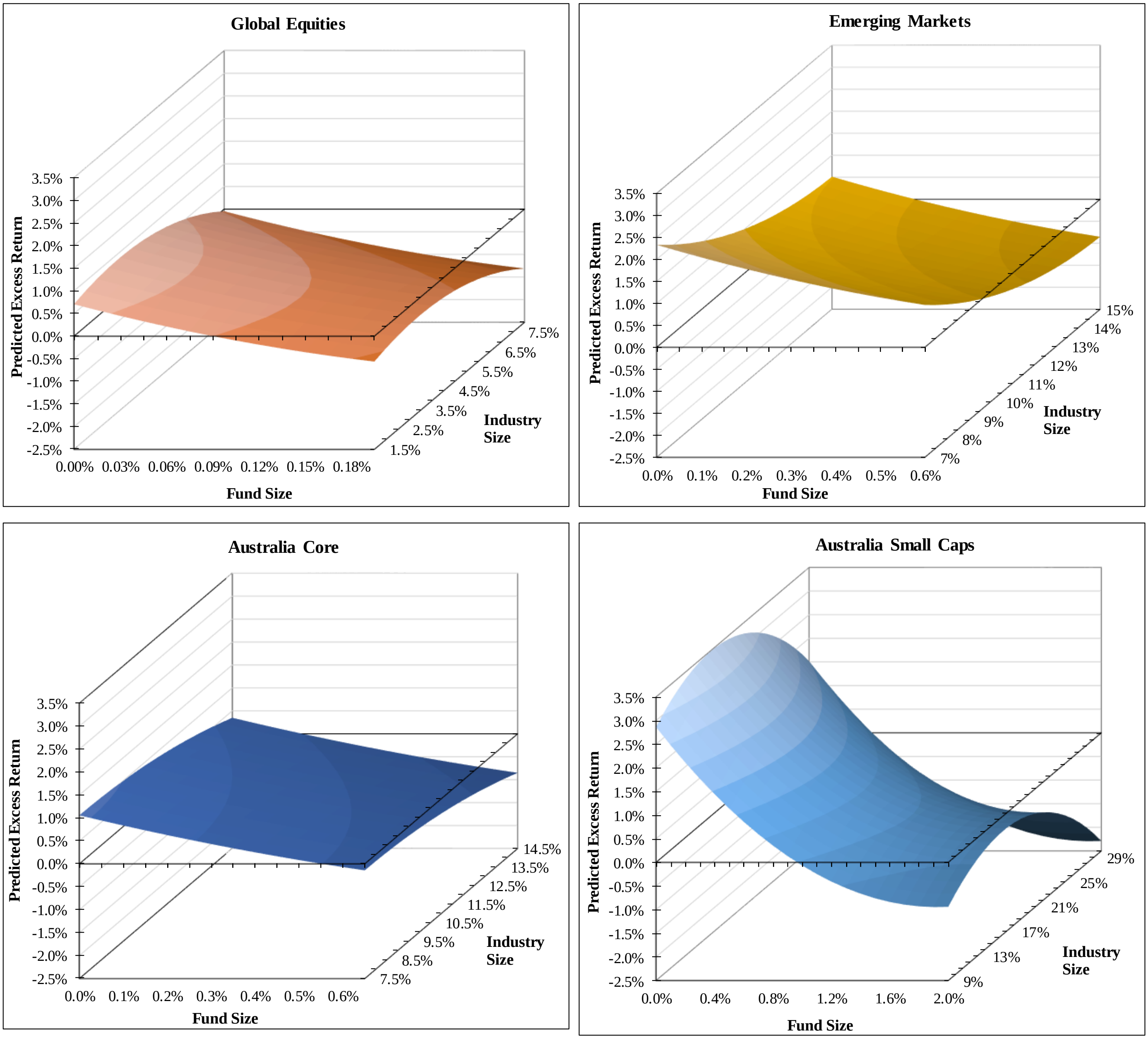


Figure A2: Percentage of Funds by Number Exceeding Capacity (i.e. $\widehat{NXR} < 0$)

Figure A2 plots the time series of the estimated percentage of funds by number operating above capacity at a high and low fee in each market, as indicated by a negative value for predicted net excess return (i.e. $\widehat{NXR}_{i,t}$). The latter is estimated using equation (6), and involves combining the slopes estimated from the linear regression as reported in Panel A of Table 2 with observations of fund size (*FS*) and industry size (*IS*) in each quarter, and then adjusting for fund-specific intercepts and the assumed fee.