

Artiga González, Tanja; Dyakov, Teodor; Inhoffen, Justus; Wipplinger, Evert

**Working Paper**

## Crowding of international mutual funds

DIW Discussion Papers, No. 1937

**Provided in Cooperation with:**

German Institute for Economic Research (DIW Berlin)

*Suggested Citation:* Artiga González, Tanja; Dyakov, Teodor; Inhoffen, Justus; Wipplinger, Evert (2021) : Crowding of international mutual funds, DIW Discussion Papers, No. 1937, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<https://hdl.handle.net/10419/233048>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# 1937

## Discussion Papers

Deutsches Institut für Wirtschaftsforschung

2021

# Crowding of International Mutual Funds

Tanja Artiga Gonzalez, Teodor Dyakov, Justus Inhoffen and Evert Wipplinger

Opinions expressed in this paper are those of the author(s) and do not necessarily reflect views of the institute.

#### IMPRESSUM

© DIW Berlin, 2021

DIW Berlin  
German Institute for Economic Research  
Mohrenstr. 58  
10117 Berlin

Tel. +49 (30) 897 89-0  
Fax +49 (30) 897 89-200  
<http://www.diw.de>

ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:  
<http://www.diw.de/discussionpapers>

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:  
<http://ideas.repec.org/s/diw/diwwpp.html>  
<http://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html>

# Crowding of International Mutual Funds <sup>\*</sup>

Tanja Artiga Gonzalez<sup>†</sup>, Teodor Dyakov<sup>‡</sup>

Justus Inhoffen<sup>§</sup> and Evert Wipplinger<sup>¶</sup>

2021-03-13

## Abstract

We construct a fund-specific measure of crowding using the equity holdings overlap of 17,364 global funds which are actively managed. Funds in the top decile of crowding underperform the passive benchmark by 1.4% per year. The poor performance cannot be attributed to fees and transaction costs alone. When we explore the economics behind crowding, we establish that the diseconomies of crowding are distinct from the ones associated with size. Among several possible mechanisms, we find support for a) a preference for liquid stocks among crowded funds and b) differences in the propagation of price pressure from flows of connected funds. Our findings reveal that the tendency of managers to follow correlated strategies is a major source of diseconomies in the active fund industry.

**Keywords:** Mutual funds, crowding, diseconomies of scale

**JEL Classification:** G15, G23

---

<sup>\*</sup>We would like to thank participants at the 2020 Annual Conference of the German Economic Association, the 3rd Dutch Network Economics Conference and seminar participants at EDHEC Business School, ESCP, VU Amsterdam, and the DIW Macroeconomics and Financial Markets for helpful comments and suggestions.

<sup>†</sup>VU Amsterdam, School of Business and Economics (SBE), De Boelelaan 1105, 1081 HV Amsterdam, the Netherlands; Tinbergen Institute. E-mail: t.artigagonzalez@vu.nl.

<sup>‡</sup>VU Amsterdam, Tinbergen Institute. E-mail: t.dyakov@vu.nl.

<sup>§</sup>DIW Berlin, Mohrenstraße 58, 10117 Berlin, Germany; VU Amsterdam. E-mail: jinhoffen@diw.de. The author's contribution was carried out in the context of the SAW-project "European capital markets and macroeconomic stability: The role of equity and debt." He gratefully acknowledges funding from the Leibniz Competition.

<sup>¶</sup>VU Amsterdam. E-mail: evert.wipplinger@vu.nl.

# 1 Introduction

The actively managed fund industry has been steadily growing over time. Funds have increased their holdings in the U.S. equity market from 2.4% in 1979 to more than 15% in the most recent decade (Pastor et al., 2015). Globally, nearly 30 trillion USD are currently managed by active funds.<sup>1</sup> With limited investment opportunities, an increasingly larger amount of capital is likely to follow correlated strategies. This in turn leads to more overlapping equity positions, an effect that we term *crowding*. In a competitive market characterized with diseconomies either on the fund or industry level (Berk and Green, 2004; Pastor and Stambaugh, 2012), an increasing amount of active capital chasing the same investment opportunities eliminates predictability in stock returns and eventually drives fund performance to zero.

In this paper, we empirically examine the relationship between crowding and fund performance. Our sample comprises 17,364 global equity mutual funds which are actively managed. In order to measure crowding, we need to establish similarities across investments. This is not a trivial task, for the following reasons. First, actual investment styles differ from the ones stated in fund prospectuses for as many as one third of all funds (Sensoy, 2009). Second, classifying funds together in the same style is arbitrary and can miss important characteristics of investment strategies. For instance, a large-cap tech fund and a large-cap precious metals fund are unlikely to hold any common positions, but both could be classified as large cap funds. Third, style classifications can vastly differ across the main data providers Morningstar, CRSP, and Thomson-Reuters. It is not a priori clear which classification is to be preferred.

We avoid the misclassification errors and idiosyncratic choices related to fund styles and infer crowding directly by comparing the portfolio holdings of funds on a stock-by-stock basis. For each fund, we compute crowding as the sum of portfolio holdings overlap across all funds with which it shares common equity positions. Defined this way, crowding is increasing in the number of connected funds and the magnitude of the portfolio holdings overlap. Effectively, we make inferences about crowding that stem from the resulting portfolio holdings, rather than the investment objective.

---

<sup>1</sup>According to data from the Global Asset Management report by the Boston Consulting Group for 2019.

Our main finding is a strong, negative association between crowding and subsequent performance. When funds operate in a crowded space, their performance deteriorates beyond the zero equilibrium alpha proposed by Berk and Green (2004) and becomes negative. In order to compute net alpha, we extend the approach of Berk and van Binsbergen (2015) and match every fund to an alternative investment opportunity set comprising of domestic and international index funds. We find that funds in the top decile of crowding have an alpha of -0.114% per month ( $t=-4.54$ ). The spread in performance between funds in the least and most crowded environment is -0.215% per month ( $t=-3.31$ ). We find qualitatively similar findings when we measure performance using dollar value added (Berk and van Binsbergen, 2015), Fama-French factors, or value-weighting of the funds in the portfolios. Our results cannot be explained by the impact of fees and transaction costs as funds in the top decile of crowding have a negative gross return after adjusting for exposure to common stock characteristics as in Dyakov and Wipplinger (2018) who extend the approach of Daniel et al. (1997) to international stocks.

As funds grow larger, their performance is likely to deteriorate due to diseconomies of scale: they run out of ideas or suffer from the increasing price impact of their trades (Berk and Green, 2004). At the same time, funds increasing in size are likely to hold more stocks, and thus have higher portfolio holdings overlap with other funds. This raises the question whether our crowding measure is just capturing a fund size effect. We show, however, that crowding is an economically distinct phenomenon from fund-level diseconomies of scale. Empirically, we analyze the nature of returns to scale and use both size and crowding to explain fund performance. We apply the instrumental-variables approach developed by Pastor et al. (2015) with the modifications of Zhu (2018), and find strong evidence that crowding drives fund performance, beyond size. For instance, when we measure alpha using index funds, we find that a 1% increase in assets held by competitors in the same stocks is associated with a 24bp decrease in monthly fund alpha. Moreover, crowding is a significant predictor of performance in Fama-Macbeth regressions, where we control for fund and domicile level variables. In summary, we find robust evidence that crowding has an adverse effect on fund performance that cannot be explained by fund-level diseconomies of scale.

When funds operate in a crowded space, they quickly compete away all positive alpha

opportunities. However, even large funds may find enough stock picking opportunities when they do not face a lot of competition. A relatively large fund operating in an investment environment without many competitors can generate higher risk-adjusted returns than a smaller fund operating in a crowded space. Our results support this intuition. We double sort funds on size and crowding and examine their subsequent performance. Among all size deciles, performance is decreasing in crowding. For instance, funds in the largest decile of size have a net alpha of 0.104% per month ( $t=3.720$ ) in the lowest tercile of crowding, and an alpha of -0.074 per month ( $t=-3.507$ ) in the top tercile. Thus, size adversely affects performance when funds operate among many competitors.

Why is performance decreasing in crowding? In the framework of Berk and Green (2004), an increase in the amount of capital chasing limited alpha results in zero net performance. If funds run out of ideas on which stocks to buy, they may simply hold the market and avoid incurring losses. However, the negative performance of funds in the top decile of crowding suggests that there is an additional mechanism which constrains their performance.

We explore several possible explanations. The first one is related to fund flows. As Coval and Stafford (2007) and Lou (2012) show, fund flows result in scaling of positions that create predictability in both stock and fund returns. At the same time, individual fund flows impact connected funds via the performance of their common holdings. Suppose the manager of fund A scales her positions in response to outflows. The price pressure on the underlying stocks will affect the performance of connected funds, causing outflows due to well-documented flow-performance relationship (e.g. Ippolito, 1992; Chevalier and Ellison, 1997). This results in further sales, leading to outflows in both fund A and its connected funds. Thus, fund flows can create externalities that can propagate shocks across the network.

In order to quantify flow effects, we compute fund-specific *PeerFlow* as the dollar flow of all connected funds weighted by the portfolio holdings overlap. We find some evidence of network externalities among funds most impacted by crowding. Specifically, we regress fund returns on lagged *PeerFlow* and its interaction with an indicator variable for being among the 30% most crowded funds. The estimated coefficient on the interaction is significant in explaining gross alpha and DGTW returns, but not for net alpha. Our holdings data is on

quarterly frequency, and thus any predictive regression assumes a significant time-lag in the flow propagation mechanism. An alternative would be to explain quarterly fund returns with fund flows occurring during the same quarter. The problem with this approach is the apparent endogeneity of fund returns and *PeerFlow* as a fund’s return impacts both the contemporaneous performance and flows of its connected peers. Blocher (2016) proposes a solution to this problem and uses lagged *PeerFlow* as an instrument. We follow his approach, and find similar results. The interaction of the instrumented *PeerFlow* with the indicator variable is positively related to contemporaneous fund returns when we use gross alpha and DGTW returns. The net alpha results remain insignificant. The difference in results between net and gross returns indicate that crowded funds manage fees in order to mitigate the *PeerFlow* pressure on net performance. This is in line with the predictions of Pastor et al. (2020) who show that funds offset high liquidity costs by charging lower fees.

In Berk and Green (2004), size erodes performance as large funds incur disproportionately larger costs. The equilibrium response of such funds is to increase holdings in liquid stocks (Pastor et al., 2020). A similar mechanism could be employed by crowded funds, even if they are not large in size. If crowded funds are concerned about the potential liquidity costs following unexpected future outflows, they increase their holdings in liquid stocks. This leads to relatively lower expected returns in comparison to funds operating in less crowded environments, who are more inclined to hold illiquid stocks and earn the associated premium. Our results offer support for this mechanism. We study funds’ trading and find that the demand for liquid stocks is monotonically increasing in crowding. The liquidity pattern we document is also consistent with Colla and Mele (2009) who show that when traders have correlated signals, their profits decrease while the liquidity of the stocks they crowd into improves. In addition, the liquidity factor of Pastor and Stambaugh (2003) can explain about a quarter of the spread in performance between funds in the top and bottom decile of crowding. Thus, preference for liquid stocks can at least partially explain the effect of crowding on performance.

Funds could crowd into the market portfolio, as they have a preference for liquid stocks. Alternatively, they could buy the market portfolio simply because they run out of investment ideas. Both arguments raise the question whether proximity to the market portfolio is the



major driver of the negative relationship between crowding and performance. To investigate this possibility, we double sort funds along two dimensions, their active share (Cremers and Petajisto, 2009) with respect to the MSCI World index and crowding. We find that crowding can explain returns over and above active share and that the impact of crowding is concentrated among funds with low active share. In addition, we find that the high concentration of MSCI stocks in the portfolios of crowded funds is driven by overweighting U.S. equity. The average weight of the U.S. is 56% in the MSCI World Index during our sample period versus 80% among the most crowded funds. These findings indicate that rather than scaling the global market portfolio, funds crowd into U.S. equity. Therefore, we conclude that our crowding measure has information content beyond deviations from the market portfolio.

The theory of Stein (2009) offers another possible mechanism for explaining negative returns to crowding. He considers sophisticated investors whose demand for a given asset is an increasing function in asset returns but who do not have an estimate of the fundamental value of the asset. Initial trading exacerbates the return signal, and investors cannot determine how much active capital is already employed by other market participants. Because of this coordination problem, investors push prices away from fundamentals and thus inflict negative externalities on each other. Momentum trading is a natural candidate to investigate these implications. It represents the most prominent investment strategy based on a return signal, and its performance is hard to explain with rational models (Asness et al., 2013). We find that the demand for momentum stocks is monotonically increasing in crowding. However, we find no relationship between momentum trading and subsequent performance. Among funds operating in a crowded space, we find no differences in subsequent fund performance between periods of heavy and low momentum trading. Thus, our findings cannot lend support to the coordination problem outlined above. We acknowledge, however, that these findings could be attributed to the low power of our test as the identification comes from the time-series.

Our sample consists of relatively more non-U.S. domiciled funds. Thus, another potential explanation why crowded funds destroy value is that foreign domiciled funds in our sample suffer from an informational disadvantage among U.S. equity while investments in U.S. stocks are overrepresented among crowded funds. The negative returns of crowded funds could be

driven by foreign funds with poor stock selection abilities among U.S. equity. To investigate this, we split our sample along fund domicile and find that the relationship between crowding and performance remains for U.S. domiciled funds. Thus, an informational gap between domestic and foreign funds is unlikely to drive our findings.

In addition to the direct effect of common stock ownership, crowding could propagate from funds that are not directly connected.<sup>2</sup> We compute an alternative measure of crowding, based on the eigenvalue centrality of funds in the global network of overlapping holdings. Specifically, we allocate higher scores to funds that are linked (i.e., have a higher portfolio overlap) with funds which in turn have relatively stronger links to other funds, which in turn have stronger links to other funds, and so on. Results are qualitatively similar.

We contribute to the literature on capacity constraints in the active fund industry. A number of studies show that fund or industry size erode performance (e.g. Chen et al., 2004, Pastor et al., 2015, Zhu, 2018). Our results show that a major source of diseconomies stems not from the fund’s size itself, but rather from the amount of capital employed by other sophisticated investors in correlated strategies. After we control for crowding, performance does not seem to decrease with size. In fact, large funds can generate positive alpha when they operate among few competitors.

Our work is also related to the literature on competition among mutual funds. Hoberg et al. (2017) study mutual funds that compete with a narrowly defined set of rivals, based on investments in similar styles. They find that funds with fewer rivals perform better than funds with many competitors. Our focus is different. We investigate the negative effects of competition and the subsequent negative performance of funds operating in a crowded space. In a robustness test, we show that the impact of crowding on fund performance remains statistically significant after controlling for the two competition measures used by Hoberg et al. (2017). Our measure of crowding is closer to the measure of competition used by Wahal and Wang (2011). They study the impact of entrants on incumbent funds and find that the active fund industry is characterized

---

<sup>2</sup> For instance, a value-investing fund A may compete with fund B, who specializes in both value stocks and small stocks. Fund C – a fund specializing in small stocks, may increase its presence among small stocks and increase the competitive pressure on fund B. This may in turn prompt fund B to become more active in value stocks, thus also crowding the investment environment of fund A.

with high levels of competition. We differ in that we study crowding using portfolio holdings overlap across all funds in the market, while they limit their scope to the impact of newly entering competitors.

Our findings on the impact of crowding among mutual funds stand in contrast to those of Sias et al. (2016), who find a positive relationship between crowding and subsequent performance of stocks held by hedge funds. A possible explanation for this difference is that crowding among mutual funds is concentrated in liquid stocks with lower levels of information asymmetries, while hedge funds are more likely to crowd into stocks that mutual funds underweight. In addition, crowding can be related to herding. A key difference is that our measure of crowding is based on aggregate portfolio overlap, while the herding literature is primarily concerned with the correlation of portfolio changes across different funds. Early studies, such as Lakonishok et al. (1992), Grinblatt et al. (1995) and Sias (2004), find that herding among mutual funds adds value. However, more recent work by Dasgupta et al. (2011) and Jiang and Verardo (2018) find that herding leads to negative future returns. Our results are consistent with the latter.

## **2 Sample construction and risk-adjustment**

### **2.1 Sample construction**

Our sample is based on quarterly positions of international mutual funds from Factset and stock level information from Datastream and Worldscope between 2001 and 2014. In addition, we collect data on fund net returns and total net assets from Morningstar Direct. Factset collects the positions of more than 90,000 funds domiciled in 89 countries and covers active and passive mutual funds, insurances, pension funds, and other funds. The data covers both alive and defunct funds. Following earlier studies (e.g., Chuprinn et al., 2015), we exclude fund reports before 2001 because coverage of Factset prior to this year is limited. We also exclude funds with net assets of less than 15 million USD due to potentially biased data (see Elton et al., 2001 and Chen et al., 2004). We drop funds not classified as either open-ended or offshore. Including offshore funds in our sample ensure that we cover funds from Europe’s two most prominent domiciles – Luxembourg and Ireland. We further keep only funds that hold at least 50 stock

holdings (equities and/or depository receipts) in their portfolios. This way we capture only funds with active equity components and exclude funds that hold equities for diversification purposes only. This procedure also ensures we drop funds from countries with lax portfolio reporting regulations, such as Australia, where funds are required to report only their top 10 portfolio holdings. We match the reported fund holdings with stock specific information from Worldscope and Datastream using CUSIP, ISIN and SEDOL identifiers. We match Factset to Morningstar using ISIN. Additional data cleaning procedures are provided in Appendix A.

We provide descriptive statistics of our sample of funds and their portfolio weights by regions in Table 1. In total, our sample covers 17,364 unique active equity mutual funds with average total assets under management (TNA) of 592 million USD. Similarly to Khorana et al. (2005), we find that there are more European than North-American funds, due to the large number of funds domiciled in the offshore locations of Luxembourg and Ireland. However, international funds are on average much smaller in size than North-American funds. On average, funds have a net alpha of -0.04% per month. Funds from each region exhibit a home bias and investing relatively more in their respective region's equity than funds domiciled elsewhere.

## 2.2 Risk adjustment

In order to adjust for risk, we compare the performance of each fund with a set of alternative investment opportunities as represented by low-cost passive funds (Berk and van Binsbergen, 2015). This approach has a number of advantages with respect to the traditional factor based models. First, factor portfolios, such as the Fama-French factors, do not incorporate transaction costs, trade impact, and trading restrictions (Huij and Verbeek, 2009). Hence, they are unlikely to represent the true alternative investment opportunity set. This problem is exacerbated within international markets, as investors throughout our sample do not have the opportunity to invest in e.g. momentum funds.

For each fund  $i$ , net alpha in month  $t$  is defined as the fund's return minus the return on the set of passive funds:

$$\alpha_{i,t} = R_{i,t} - \sum_{j=1}^{n(t)} \beta_f^b R_t^b, \quad (1)$$

where  $R_{i,t}$  denotes the net excess return of fund  $i$  in month  $t$ ,  $R_t^b$  is the excess net return earned by investors on the  $n$ th index fund at time  $t$ , and  $\beta_f^b$  is the sensitivity of fund  $i$  to the  $b$ -th index fund. As reflected in the notation, the number of available benchmark funds may vary over time. When we compute gross alpha, we use gross fund and index fund returns, instead. We source net returns from Morningstar Direct and compute gross returns from the last published portfolio holdings in Factset. We require at least 48 monthly observations in order to compute alphas. Finally, we employ the methodology of Berk and van Binsbergen (2015) to compute dollar value added as net alpha times total net assets.

A challenge in applying the above approach lies in the selection of passive funds. To avoid a bias in selecting index funds, we follow Berk and van Binsbergen (2015) who select Vanguard index funds as benchmarks for U.S.-based funds. The selection of Vanguard index funds for investments in non-U.S. markets follows Dyakov et al. (2019), who extend Berk and van Binsbergen (2015) to international markets. We check if a fund invests on average more than 75% in one of the four broad geographical regions that we define alongside MSCI country classification: i) developed North-America (NAM), ii) developed Europe (EUR), iii) developed Asia-Pacific and Japan (APA), and iv) Emerging Markets (EME). If the condition is met, we allocate the fund to that region. If not, we refer to the fund as a global (GLO) fund. Next, we specify the set of passive Vanguard for each region and use equation (1) to compute the fund's net alpha. The selection of Vanguard funds for each region is described in Appendix B.

In addition, we measure performance by comparing the return of every stock  $k$  with a set of stocks with similar size, book-to-market, and momentum characteristics (also known as DGTW adjusted returns, following Daniel et al., 1997, Wermers, 1999, and Wermers, 2003, who introduced this methodology to U.S. stocks). Specifically, the characteristic-adjusted return on a stock  $k$  is given by

$$\alpha_{k,t}^{DGTW} = R_{k,t} - R_{k,t}^{bench}, \quad (2)$$

where  $R_{i,t}^{bench}$  denotes the return of a benchmark portfolio of stocks with similar size, book-to-market, and momentum characteristics. Fund  $i$ 's return is then simply the weighted average of

the  $\alpha^{DGTW}$  across all stocks it holds.<sup>3</sup> The DGTW returns may not capture the true alternative investment opportunity set, as managers might be constrained in the stocks they could buy, due to trading costs, regulations, or other frictions. However, they offer a direct risk-adjustment as each stock is matched to a benchmark of stocks most similar in the size, book-to-market, and momentum space. In addition, calculated alphas are not affected by estimation error. Lastly, DGTW returns are computed in local currency and are hence not affected by swings in exchange rates. Consistent with these arguments, Dyakov et al. (2019) find that DGTW returns can better detect the impact of capacity constraints on performance among international funds.

### 3 Crowding

#### 3.1 Construction of Crowding

Our main crowding measure is based on portfolio holdings overlap. It is computed in two steps. First, for each pair of funds, we compute portfolio similarity as the weight in common stocks. Second, for each fund at each point in time, we aggregate its portfolio similarity with all other funds in the sample.

Specifically, for any two funds  $i$  and  $j$  in our sample, we compute the common portfolio holding  $e_{ij}$  as the sum of the minimum portfolio weight in a stock across all assets. The higher the overlap of portfolio holdings of two funds, the stronger the link of the two funds in our network. Let  $\omega_i^k$  and  $\omega_j^k$  denote the weights of fund  $i$  and fund  $j$  in stock  $k$ . The total portfolio overlap between fund  $i$  and  $j$  is then defined as

$$e_{ij} = \sum_{k \in P_i \cap P_j} \min(\omega_i^k, \omega_j^k) \quad (3)$$

where  $P_i$  is the set of stocks fund  $i$  is invested in. This measure of overlap is symmetric (in network terms, the connection is undirected with  $e_{ij} = e_{ji}$ ) and ranges from 0 to 1. If two funds hold the exact same portfolio of stocks and in the same proportion, then the overlap measure is

---

<sup>3</sup>See Appendix E in Dyakov et al. (2019) who provide a detailed methodology for computing benchmark-adjusted returns for international stocks belonging to broad geographical regions and the solutions for problems associated with differences in the size of equity markets and accounting standards.

one. Alternatively, if the pair of funds does not hold any asset in common, the measure equals zero.

Articles that analyze networks of common asset holdings have used a number of other measures, such as common ownership (Asness et al., 2013), the Jaccard Index (e.g. Hüser and Kok, 2019), cosine similarity and absolute differences in portfolio weights (e.g. Sias et al., 2016). We use common weight due to the simplicity of the measure. Our overlap measure has one notable drawback that it shares with the measures based on individual stock weights mentioned above. It does not recognize that stocks with similar characteristics are potential substitutes. For instance, two large growth funds may have zero overlap portfolio overlap, while holding essentially stocks with the same profile.

After having computed the portfolio overlap  $e_{ij}$  between two funds, we derive the fund-specific crowding score by summing up the pairwise overlaps with all other funds in the universe. For each fund  $i$  the crowding score is given by

$$crowd_i = \sum_{j \in Q, j \neq i} e_{ij} \quad (4)$$

with  $Q$  being the universe of mutual funds observed in the respective quarter of our sample. The crowding score measures how similar the portfolio holdings of a mutual fund are to the portfolio holdings of all other funds in the network. The higher the crowding score for a fund, the more it competes with other funds for the same investment opportunities. Wahal and Wang (2011) use a similar measure to study the impact of newly entering funds on incumbent funds. While they construct portfolio overlap in a similar way, they only sum across edges with new entrants.

### 3.2 Descriptive Statistics

Table 2 presents fund characteristics for fund portfolios sorted into deciles based on crowding. The average total net assets (TNA) of funds in the lowest crowding decile is 302 million USD while the average TNA of funds in the top decile is nearly three times as large. Crowded funds are more diversified across firms, countries, and industries. They are also characterized with

a significantly higher inverse of the Herfindahl Index of industry share, thus also indicating higher diversification. Panel B shows average portfolio weights for stock regions. Funds with a high crowding score invest more capital into North American stocks while funds with a lower score invest more prominently in Emerging Markets, Frontier Markets, and the Asia Pacific and Japan region.

In addition, we provide stock-level characteristics in Table 3. Pastor et al. (2020) predict that larger funds hold more liquid assets and are more diversified. Their predictions extend to crowded funds, because crowding captures part of the size effect. In line with their prediction, we find that crowded funds have a preference for liquid stocks, as proxied for by Amihud illiquidity (Amihud, 2002). In addition, they hold less risky assets characterized with larger size, lower book-to-market, and lower momentum and volatility as well as stocks with lower levels of information asymmetries, as proxied for by the higher number of analysts covering them. Furthermore, funds operating in a crowded space hold more mature firms and stocks with higher dividend yield. Lastly, they diversify relatively more across foreign stocks and geographically more distant stocks.

## 4 Crowding and Fund Performance

### 4.1 Single sorts

At the end of each quarter, we sort funds into ten deciles based on their crowding score. Next, we track the average performance of funds in the decile portfolios over the subsequent three months, and then rebalance. Time-series averages of the returns of the decile portfolios are presented in Table 4.

We find a strong, negative relationship between crowding and subsequent performance. The pattern is consistent across different risk-adjustment methods. In addition, performance for funds in the most crowded portfolios is negative. For instance, funds in the top decile of crowding exhibit a net alpha of -0.114% per month ( $t=-4.54$ ). The spread in net alpha between funds in the most and the least crowded environment is -0.215 per month ( $t=-3.31$ ). Funds in the top decile of crowding feature a negative amount of dollar value added: -1.855\$ million



per month. Note that fees and transaction costs alone cannot explain the negative overall performance of the most crowded funds, as the return patterns are similar when we use gross instead of net returns. In addition, the results are robust to using value-weighting of funds in the decile portfolios, as well as to using various traditional factor model specifications. The findings are presented in the Appendix, Table A5.

## 4.2 Crowding and Diseconomies of Scale

Previous literature (e.g. Chen et al., 2004) shows a similar monotonically decreasing pattern of fund performance when sorting on fund size. Large funds typically operate in a more crowded environment. Therefore, crowding is likely to capture at least some of the effects of fund size. In this Section, we empirically disentangle the role of the two variables in explaining fund performance.

Consider a group of mutual funds, indexed  $i = 1, \dots, N$ . The dependent variable in our analysis is  $r_{i,t}$ , which denotes the risk-adjusted return (alpha) of fund  $i$  in month  $t$ . The total market value of the fund at the end of the previous month is  $q_{i,t-1}$ . We want to estimate the effect of lagged crowding  $crowd_{i,t-1}$  and lagged size  $q_{i,t-1}$  on fund performance  $r_{i,t}$ . Following Zhu (2018), crowding and size are expressed in logs due the wide range of their values:

$$r_{i,t} = a_i + b_1 \log crowd_{i,t-1} + b_2 \log q_{i,t-1} + \varepsilon_{it}. \quad (5)$$

In this equation,  $a_i$  are fund fixed effects, absorbing the cross-sectional variation in managerial skills, which are assumed to be time-invariant.<sup>4</sup> The coefficient  $b_1 < 0$  identifies the adverse effect of crowding on performance, while the coefficient  $b_2 < 0$  identifies decreasing returns to scale at the fund level.

A standard fixed effects estimator requires the regressors in equation (5) to be strictly exogenous. That is, regressors should be uncorrelated with  $\varepsilon_{it}$  across all time periods. For equation (5), however, this is not the case as (a) fund size mechanically relates to past performance (even without flows), and (b) investor flows respond to past performance (Pastor et al.,

---

<sup>4</sup>In general, this term would also capture fund-specific “skill” related to, for example, operational costs at the fund or family level.

2015). To address this problem, we follow the methodology outlined in Pastor et al. (2015) and Zhu (2018). We forward demean all variables and follow a two stage least squares approach (2SLS). In a first stage regression, we use backward-demeaned fund size and lagged fund size as instruments for forward-demeaned size. In the second state, we regress forward demeaned alpha on the fitted values from the first stage as well as forward-demeaned crowding. We elaborate on this methodology in Appendix F.

In addition to  $crowd_i$ , we introduce a new variable,  $PeerTNA_i$ , which sums the product of the total portfolio overlap between fund  $i$  and  $j$  with the size of fund  $j$ , across all funds that overlap with  $i$ :

$$PeerTNA_{i,t-1} = \sum_{j \in Q} e_{ij,t-1} q_{j,t-1} \quad (6)$$

Both  $crowd_i$  and  $PeerTNA_i$  quantify how crowded the investment environment of fund  $i$  is. However, the advantage of  $PeerTNA_i$  is that it is a dollar measure, which allows us to better gauge the economic impact of crowding on performance. Thus, in additional specifications, we run the 2SLS model with  $\log PeerTNA_{i,t-1}$  in place of  $\log q_{i,t-1}$ .

The results are reported in Table 5. In Panel A, we use net alpha as performance measure. Our crowding measures are significant in all specifications. After controlling for fund size, the estimated coefficient on  $crowd_i$  is -0.0020 ( $t=-2.68$ ). This is an economically large effect, as evidenced by the estimated coefficients on  $PeerTNA_i$ . A one percent increase in the assets managed by competitors translates into a 21bp drop in performance ( $t=-2.43$ ).  $crowd_i$  remains significant in specifications where we use gross alpha and DGTW returns as performance measures. In Appendix E, we examine the robustness of our findings and find that crowding explains performance in Fama-Macbeth regressions where we include control variables and domicile fixed effects. Our results indicate that crowding is an economically distinct phenomenon from the previously documented diseconomies associated with fund size.

In Table 5, we find limited evidence about the role of fund size in explaining performance. The estimated coefficients are marginally significant only for gross alpha. These findings are consistent with Dyakov et al. (2019), who study diseconomies of scale among geographical regions

of investment. They find that diseconomies of scale are weaker outside of the U.S., possibly because the active industry is below its optimal size. The lower levels of active investment indicate that capacity constraints are less binding, and thus more difficult to estimate in the data.

We offer additional insights about the role of crowding and size as determinants of performance. On the one hand, a large amount of funds chasing limited investment opportunities means that alpha quickly disappears. Even a skilled manager operating a small fund may find difficulties in beating the benchmark. On the other hand, if funds operate in a less crowded environment, mispriced stocks might be easier to identify. When there is little competition, even large funds may be able to generate superior performance. Thus, the effect of size on performance depends on how crowded the investment environment is. Our results support this argument. We double sort funds into decile portfolios of size and then into crowding terciles, and examine their subsequent performance. Results are reported in Table 6. Among the largest funds, funds operating in the lowest tercile of crowding are able to beat the benchmark. Their average net alpha is 0.104% per month ( $t=3.720$ ). In contrast, funds in the top tercile of crowding have an average net alpha of -0.074 per month ( $t=-3.507$ ). Among all size deciles, funds operating in a crowded environment generate significantly lower returns.

## 5 Costs of Crowding

Our main result of decreasing performance with increasing crowdedness is consistent with the predictions of Berk and Green (2004) where performance decreases with scale. In their model, active fund managers cannot infinitely scale their investment opportunities. Either the price impact of their trades increases, or they eventually run out of ideas. In equilibrium, investors reward individual fund managers with capital up to the point where returns going forward are zero. That is, net alpha must be zero and gross alpha equals the fund's costs. In our context, this implies that the returns of crowded funds should be nonnegative, before fees and transaction costs. While managers could invest excess capital into crowded stocks with negative performance, they should only be able to raise additional excess capital while the

fund’s overall performance remains positive or zero. However, we find a negative performance of the most crowded funds in Table 4. This indicates that additional frictions or unaccounted risk premiums are associated with crowding, driving aggregate performance negative. In this Section, we explore several potential mechanisms.

## 5.1 Peer Flows

It is well known that mutual fund flows can have a significant impact on subsequent fund returns (Coval and Stafford, 2007; Lou, 2012). Hence, we investigate the role of fund flows in creating externalities (Blocher, 2016; Anton and Polk, 2020) among crowded funds.

In the context of this paper, fund flows can be considered as individual fund shocks that can propagate among the network of funds connected via their common portfolio holdings. For instance suppose all funds have investors that are chasing returns. When Fund A faces poor performance, investors will withdraw capital. Further suppose Fund A scales its portfolio as it sells assets to meet redemptions. Fund B holds some stocks in common with Fund A and will exhibit a lower performance in these stocks due to the price pressure from Fund A. Since the investors of Fund B also chase returns, they will tend to withdraw capital from Fund B in the short-term as well. This process would be repeated among all connected funds and creates a feedback loop. In turn, this mechanism creates additional costs or externalities as funds are forced to trade not only based on their investment strategy but also in response to flows induced by the performance of other funds.

Our focus is on the effect of flows among funds with direct connections. To this end, we construct a measure of neighboring peer flows  $PeerFlow$  for each fund  $i$  and quarter  $t$  as the sum of the product of portfolio overlap  $e_{i,j}$  and flows of connected funds  $Flow_{j,t}$ :

$$PeerFlow_{i,t} = \sum_{j \neq i} e_{i,j} Flow_{j,t}. \quad (7)$$

Hence, in computing peer flows for an individual fund, flows of more similar funds receive higher weights than flows of funds with less portfolio overlap. In Panel A of Table 7, we provide estimates of the predictive power of peer flows for subsequent performance. Peer flows do not

significantly predict performance in subsequent quarters for all funds. However, we do find evidence for a significant difference in the predictive power of peer flows between crowded and non-crowded funds. Higher peer flows for the 30% most crowded funds have a more positive effect on future performance than compared to the rest of the funds. The effect is highly significant at the 1% confidence level for both measures of gross performance, gross alpha ( $t = 2.26$ ) and DGTW-returns ( $t = 3.32$ ). This holds while we control for the overall level of fees of funds, but not for net performance as the coefficient of the interaction effect is not significant when predicting net alpha.

We also explore the contemporaneous relationship as our holdings data are on a quarterly frequency which can only reveal a slow propagation mechanism. A challenge in measuring the contemporaneous impact is that fund performance affects both the contemporaneous flow of a fund as well as that of its peers. To address this endogeneity issue, we employ a two stage least squares (2SLS) methodology where we instrumentalize peer flow with lagged peer flow as in Blocher (2016). The results, presented in Panel B of Table 7, are qualitatively similar to those of the predictive regression in Panel A. Again, the key observation is that the gross performance of less crowded funds is more isolated from network effects of flows than that of the 30% most crowded funds. Note that peer flow already takes crowding into account by scaling flows with crowding in Equation (7). This suggests that the difference in flow sensitive between crowded funds is also driven by the characteristics of the stocks they choose to crowd into. We further investigate such relationship in the next sections.

## 5.2 Preference for Liquidity

Pastor et al. (2020) provide evidence of diseconomies of scale which is based on trade offs between funds' choices of characteristics. In equilibrium, funds endogenously choose their portfolio and fee structure such as to offset trading costs. Their arguments extend to crowded funds. First, more crowded funds will have a higher demand for liquid stocks because they need to offset trading costs associated with their more interlinked and concentrated portfolios. Second, the prediction of Berk and Green (2004) that excess capital will be allocated to passive investments such as the market portfolio also implies a preference for liquid stocks. At the margin funds

will scale their passive investment of excess capital to meet additional flows. Therefore, funds will prefer more liquid stocks to offset the trading costs associated with flows such as the higher impact of peer flows on returns for crowded funds documented in the previous section.

Hence crowded funds should not only have a higher overall exposure to more liquid stocks as shown in Table 3 but also a higher demand for liquid stocks among their trades. Following Sias (2004), we first compute a buyer ratio  $BR$ , defined as the number of funds buying a stock  $k$  each quarter relative to the number of funds trading the stock

$$BR_{k,t} = \frac{\# \text{ of funds buying stock } k}{\# \text{ of funds buying stock } k + \# \text{ of funds selling stock } k} \quad (8)$$

We then define the demand for funds in stock  $k$  as the standardized value of  $BR$

$$ID_{k,t} = \frac{BR_{k,t} - \overline{BR_{k,t}}}{\sigma(BR_{k,t})} \quad (9)$$

where  $\overline{BR_{k,t}}$  and  $\sigma(BR_{k,t})$  stand for the cross-sectional quarterly mean and standard deviation of  $BR_{k,t}$ , respectively. The standardization of the variable allows us to compare funds' demand across time periods and crowding deciles.

Each quarter, we estimate a cross-sectional regression of demand on stock characteristics

$$ID_{k,t} = \alpha_t + \beta_t \mathbf{X}_{k,t-1} + \gamma_t ID_{k,t-1} + \epsilon_t \quad (10)$$

where  $\mathbf{X}$  is a vector of stock characteristics, including size, value, momentum and illiquidity. Following Sias (2004), we include lagged stock demand as funds tend to herd into the same stocks over adjacent quarters. Since we study stock demand conditional on crowding, we estimate equation (10) separately for funds belonging to the ten crowding deciles.

We report time-series averages of the estimated coefficients in Table 8. The coefficient on the aggregate Amihud (2002) illiquidity score is always negative and generally decreasing in crowding among fund portfolios. Starting with decile 4 up to the highest decile of crowding, the coefficient is negative at confidence levels of 5% and above. The pattern indicates higher trading demand for liquid stocks among more crowded funds.

Therefore, a possible explanation for the negative returns to crowding is a liquidity premium unaccounted for in the risk-adjustment of Section 2.2. To explore this, we compare the results of a risk adjustment using the Fama-French factors with and without an augmented liquidity factor (Pastor and Stambaugh, 2003) in Table 9. The difference in risk-adjusted performance between the least and most crowded decile of fund portfolios amounts to -0.214% per month ( $t = -2.79$ ) in the unaugmented Fama-French factor model. After we include the liquidity factor the difference in performance is about a quarter smaller at -0.166% per month ( $t = -2.27\%$ ). Moreover less crowded funds exhibit a significantly lower loading on the liquidity factor. This reveals that less crowded funds are able to earn more of a liquidity premium on their portfolio. Overall our results indicate that stock liquidity can partially but not fully explain the negative returns to crowding.

Another way to reduce trading costs in Pastor et al. (2020) is diversification because portfolio liquidity is composed of stock liquidity and portfolio diversification. We investigate this in Tables 11 and 12. First, we double sort funds on their active share (Cremers and Petajisto, 2009) with respect to the MSCI World portfolio and then crowding. Table 11 shows that crowding has explanatory power beyond deviations from the market portfolio. Crowding generates performance differences among funds with low active share. Specifically, the spread between high and low crowding among the five lowest deciles of active share ranges between -4 and -15 basis points, with significant p-values. Thus, some funds are able to stick close to the market portfolio, without having an excessive overlap with other funds that decreases their performance. On the other hand, the spread is insignificant for funds with high active share (decile portfolios 6-10). As high active share funds stray away from the benchmark, they are more likely to pick stocks where the effects of crowding are less pronounced. Next, we relate the average exposure of funds to broad MSCI geographic regions. Results are summarized in Table 12. North-American stocks represent 55% of the MSCI World Index, but 80% among the top 5% most crowding funds. Thus, rather than scaling the global market portfolio, funds crowd into U.S. equity.

### 5.3 Coordination Externalities

Stein (2009) describes a coordination problem that arises when sophisticated investors face incomplete information. There are two important features of his model. First, investors are unaware of the total amount of capital that other sophisticated investors have allocated towards the same strategy. That is, they are unaware of the total number of competing investors and of their specific choice of investment strategies. Second, trading is not based on a fundamental anchor. Instead, the demand for assets is driven by their prices. Stein (2009) shows that the inability of investors to condition their trades on that of others gives rise to a coordination problem: an unexpectedly large number of competing investors adopting the same strategy, yields overreaction in prices. However, in equilibrium, prices underreact as more capital is deployed in times of overreaction than in times of underreaction. Thus, investors can inflict negative externalities on each other. Managers in crowded funds are sophisticated investors that have chosen the most similar strategies to others and in this setup would constitute investors most likely to overreact to the price signal. They would tend to buy overvalued assets and generate negative future returns, once prices revert back to fundamentals.

To explore this possibility, we focus on momentum (Jegadeesh and Titman, 2002) as the most common example of a trading strategy based on price movements rather than on fundamentals. Momentum strategies are vastly popular among investors and the academic literature on momentum has failed to pinpoint a rational explanation for its success (Asness et al., 2013). Thus, the demand for momentum stocks is likely to be solely determined by the performance of past losers and winners.

First we inspect the momentum loading of the demand regression given in Equation (10). The corresponding coefficient estimates in Table 8 exhibit a positive relationship between crowding and demand for momentum stocks. The average estimated coefficient on momentum among funds is generally increasing with crowding: it is negative but statistically insignificant ( $t = -0.006$ ) in the bottom decile of crowded funds but positive and highly statistically significant ( $t = 8.54$ ) in the most crowded decile. The findings indicate that a potential coordination problem is likely to be more pronounced among crowded funds.



However, a stronger demand for momentum stocks does not necessarily indicate that the coordination problem drives the negative performance. An additional prediction is that the price impact of crowding should be more pronounced when funds display a relatively stronger demand for momentum stocks. We therefore split our time periods into 5 groups, based on the estimated quarterly coefficients on momentum from equation (10). Table 10 shows the time-series averages of the subsequent three month and 12 month returns. When funds in the top decile of crowding trade relatively more on momentum and subsequent risk-adjusted returns are negative. However, future returns are also negative, even in periods when crowded funds do not heavily trade on momentum. Thus, our findings do not lend support to the coordination problem. While crowded funds display stronger demand for momentum stocks, momentum trading is unlikely to drive the negative returns of funds operating in a crowded space.

## 6 Additional tests

We conduct additional tests, in order to explore the robustness of our findings and rule out alternative mechanisms for the relationship between crowding and performance. First, we use Fama-Macbeth regressions of returns on lagged crowding. Results are reported in Table 13. We find consistent results – crowding impacts negatively subsequent performance. In Appendix E, we report returns on portfolios sorted by crowding (or double sorted on size and crowding), where we use value-weighting of funds and more performance measures, including alphas from factor models. Again, results are consistent.

### 6.1 Informational Difference between U.S. and non-U.S. funds

Our findings indicate that crowded funds overinvest into U.S. equity. A possible explanation is that European managers are at an informational disadvantages with respect to domestic managers in U.S. equity. However, the underperformance is not driven by an informational gap between U.S. and foreign funds. When we restrict the sample to funds domiciled in the U.S. our main results hold. Specifications (1)-(4) in Table 14 show that the two crowding remain a statistically significant predictors for performance, after controlling for fund characteristics in

the sample of U.S. domiciled funds.

## 6.2 Competition and Crowding

Hoberg et al. (2017) construct fund-specific measures of competition to study performance persistence. For each fund, they define a set of competitors based on total distance in the space spanned by size, book-to-market, and momentum. According to their approach, stocks with similar characteristics are potential substitutes. On the other hand, our crowding measure is based on portfolio holdings overlap and therefore pinpoint to the individual stocks where funds crowd. The implications from our findings are different. Hoberg et al. (2017) stress performance persistence among funds facing less competitive pressure whereas we point to the negative impact of crowding on performance in general.

We find that the negative relationship between our crowding measures and subsequent performance remains after including the two measures of competition used by Hoberg et al. (2017). Closely following their work, we identify a fund-specific set of rivals based on overall similarity in the size, book-to-market, and momentum dimensions. For each fund, *TSIM* measures total similarity across all rivals, while *NPeers* measures the number of rivals. We restrict the analysis to a similar set of U.S. domiciled funds used by Hoberg et al. (2017). Specifications (5)-(8) in Table 14 present the results. In all specifications, our two crowding measures remain statistically significant.

## 7 Conclusion

Our paper is important for understanding the role of scale in the active mutual fund industry. We show that crowding is associated with diseconomies that drive performance negative. The negative effects associated with crowding are not caused by fund size. We explore a few potential mechanisms for the effect of crowding on performance. We find support for two them. The first one is a flow channel. Shocks in a crowded can transmit from one fund to another and thus generate negative externalities among funds with high portfolio holdings overlap. The second one is a preference for liquid stocks. Funds with many competitors choose to hold more liquid

stocks, as they need to be able to respond to potential negative shocks and subsequent outflows. This, in turn, constrains their performance.

## References

- Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of financial markets* 5(1), 31–56.
- Anton, M. and C. Polk (2020, November). Connected stocks. *The Journal of Finance* 69(3), 1099–1127.
- Asness, C. S., T. J. Moskowitz, and L. H. Pedersen (2013, June). Value and momentum everywhere. *The Journal of Finance* 68(3), 929–985.
- Berk, J. and R. Green (2004). Mutual Fund Flows and Performance in Rational Markets. *Journal of Political Economy* 112(6), 1269–1295.
- Berk, J. B. and J. H. van Binsbergen (2015). Measuring Skill in the Mutual Fund Industry. *Journal of Financial Economics* 118(1), 1–20.
- Blocher, J. (2016). Network externalities in mutual funds. *Journal of Financial Markets* 30, 1 – 26.
- Chen, J., H. Hong, M. Huang, and J. D. Kubik (2004). Does Fund Size Erode Mutual Fund Performance? The role of Liquidity and Organization. *The American Economic Review* 94(5), 1276–1302.
- Chevalier, J. and G. Ellison (1997). Risk Taking by Mutual Funds as a Response to Incentives. *Journal of Political Economy* 105(6), 1167–1200.
- Chuprinn, O., M. Massa, and D. Schumacher (2015). Outsourcing in the International Mutual Fund Industry: An Equilibrium View. *The Journal of Finance* 70(5), 2275–2308.
- Colla, P. and A. Mele (2009, 05). Information Linkages and Correlated Trading. *The Review of Financial Studies* 23(1), 203–246.
- Coval, J. and E. Stafford (2007). Asset Fire Sales (and Purchases) in Equity Markets. *Journal of Financial Economics* 86(2), 479–512.

- Cremers, K. J. M. and A. Petajisto (2009). How Active is Your Fund Manager? A New Measure That Predicts Performance. *The Review of Financial Studies* 22(9), 3329–3365.
- Daniel, K., M. Grinblatt, S. Titman, and R. Wermers (1997). Measuring Fund Performance with Characteristic-Based Benchmarks. *The Journal of Finance* 52(3), 1035–1058.
- Dasgupta, A., A. Prat, and M. Verardo (2011). Institutional Trade Persistence and Long-term Equity Returns. *The Journal of Finance* 66(2), 635–653.
- Dyakov, T., H. Jiang, and M. Verbeek (2019). Trade Less and Exit Overcrowded Markets: Lessons from International Mutual Funds. *Review of Finance*. rfz014.
- Dyakov, T. and E. Wipplinger (2018). Institutional Ownership and Future Stock Returns: An International Perspective. *Review of International Finance*, forthcoming.
- Elton, E. J., M. J. Gruber, and C. R. Blake (2001). A First Look at the Accuracy of the CRSP Mutual Fund Database and a Comparison of the CRSP and Morningstar Mutual Fund Databases. *The Journal of Finance* 56(6), 2415–2430.
- Fama, E. F. and K. R. French (2012). Size, Value, and Momentum in International Stock Returns. *Journal of Financial Economics* 105(3), 457–472.
- Griffin, J. M. (2002). Are the Fama and French Factors Global or Country Specific? *The Review of Financial Studies* 15(3), 783–803.
- Grinblatt, M., S. Titman, and R. Wermers (1995). Momentum Investment Strategies, Portfolio Performance, and Herding: A Study of Mutual Fund Behavior. *American Economic Review* 85(5), 1088–1105.
- Hoberg, G., N. Kumar, and N. Prabhala (2017). Mutual Fund Competition, Managerial Skill, and Alpha Persistence. *The Review of Financial Studies* 31(5), 1896–1929.
- Huij, J. and M. Verbeek (2009). On the Use of Multifactor Models to Evaluate Mutual Fund Performance. *Financial Management* 38(1), 75–102.

- Hüser, A.-C. and C. Kok (2019). Mapping bank securities across euro area sectors: comparing funding and exposure networks.
- Ince, O. S. and R. B. Porter (2006). Individual Equity Return Data From Thomson Datastream: Handle With Care! *Journal of Financial Research* 29(4), 463–479.
- Ippolito, R. A. (1992). Consumer Reaction to Measures of Poor Quality: Evidence from the Mutual Fund Industry. *Journal of Law and Economics* 35(1), 45–70.
- Jegadeesh, N. and S. Titman (2002). Cross-sectional and time-series determinants of momentum returns. *Review of Financial Studies* 15(1), 143–157.
- Jiang, H. and M. Verardo (2018). Does Herding Behavior Reveal Skill? An Analysis of Mutual Fund Performance. *The Journal of Finance* 73(5), 2229–2269.
- Khorana, A., H. Servaes, and P. Tufano (2005). Explaining the Size of the Mutual Fund Industry Across the World. *Journal of Financial Economics* 78(1), 145–185.
- Lakonishok, J., A. Shleifer, and R. W. Vishny (1992). The Impact of Institutional Trading on Stock Prices. *Journal of Financial Economics* 32(1), 23–43.
- Landau, E. (1895). Zur Relativen Wertbemessung Der Turnierresultate. *Deutsches Wochensach* 11, 366–369.
- Lou, D. (2012). A Flow-based Explanation for Return Predictability. *The Review of Financial Studies* 25(12), 3457–3489.
- Pastor, L. and R. F. Stambaugh (2003). Liquidity Risk and Expected Stock Returns. *Journal of Political Economy* 111(3), 642–685.
- Pastor, L. and R. F. Stambaugh (2012). On the Size of the Active Management Industry. *Journal of Political Economy* 120(4), 740–781.
- Pastor, L., R. F. Stambaugh, and F. T. Taylor, Lucian A. (2020). Fund tradeoffs. *Journal of Financial Economics* forthcoming.

- Pastor, L., R. F. Stambaugh, and L. A. Taylor (2015). Scale and Skill in Active Management. *Journal of Financial Economics* 116(1), 23–45.
- Porta, R. L., F. Lopez-de Silanes, A. Shleifer, and R. W. Vishny (1998). Law and finance. *Journal of political economy* 106(6), 1113–1155.
- Sarkissian, S. and M. J. Schill (2003). The overseas listing decision: New evidence of proximity preference. *The Review of Financial Studies* 17(3), 769–809.
- Schmidt, P. S., U. von Arx, A. Schrimpf, A. F. Wagner, and A. Ziegler (2011). On the Construction of Common Size, Value and Momentum Factors in International Stock Markets: A Guide with Applications. *Working Paper*.
- Sensoy, B. A. (2009). Performance evaluation and self-designated benchmark indexes in the mutual fund industry. *Journal of Financial Economics* 92(1), 25 – 39.
- Sias, R., H. J. Turtle, and B. Zykaj (2016). Hedge fund crowds and mispricing. *Management Science* 62(3), 764–784.
- Sias, R. W. (2004). Institutional Herding. *The Review of Financial Studies* 17(1), 165–206.
- Stein, J. (2009). Presidential Address: Sophisticated Investors and Market Efficiency. *The Journal of Finance* 64(4), 1517–1548.
- Wahal, S. and A. Y. Wang (2011). Competition among mutual funds. *Journal of Financial Economics* 99(1), 40 – 59.
- Wermers, R. (1999). Mutual Fund Herding and the Impact on Stock Prices. *The Journal of Finance* 54(2), 581–622.
- Wermers, R. (2003). Is Money Really 'Smart'? New Evidence on the Relation Between Mutual Fund Flows, Manager Behavior, and Performance Persistence. *Working paper*.
- Zhu, M. (2018). Informative Fund Size, Managerial Skill and Investor Rationality. *Journal of Financial Economics* 130(1), 114–134.

**Table 1:** Descriptive Statistics

This table provides summary statistics of fund stock holdings by funds' region of domicile. We report the number of unique funds, mean total net assets under management (TNA, in million USD), mean number of stocks in portfolios, mean net alpha (in %) and the mean portfolio weights for each stock region (in %). We first compute cross-section and subsequently quarterly means. The imposed regions follow the MSCI market classification: NAM = North America, EUR = Europe, APA = Asia Pacific excluding Japan, JPN = Japan, EM = Emerging Markets, FM = Frontier Market. See table A1 in the appendix for summary statistics by funds' country of domicile.

Fund region	Funds	TNA	Holdings	Net Alpha	Stock region (average weight in %)					
	#	mill USD	mean #	mean % p.m.	NAM	EUR	APA	JPN	EM	FM
NAM	6,487	1,176	170	−0.05	78.2	10.9	3.2	4.0	2.9	0.9
EUR	9,843	277	140	−0.03	32.5	44.3	6.3	9.7	6.2	0.9
APA	138	265	97	0.20	16.8	11.0	29.8	21.0	21.0	0.4
JPN	84	69	128	−0.60	20.4	16.4	4.7	52.7	5.3	0.5
EM	519	93	71	−0.11	9.4	8.7	4.8	1.4	74.8	0.8
FM	293	116	123	0.12	30.9	27.3	9.2	10.2	8.2	14.3
All domiciles	17,364	592	153	−0.04	54.3	27.3	5.0	6.8	5.6	1.0



**Table 2:** Summary statistics of fund characteristics for decile portfolios sorted on crowding

This table provides summary statistics of fund portfolios conditional on crowding. At the end of each quarter, we sort funds into ten portfolios. We compute the cross-section average of characteristics conditional on the decile. We report fund characteristics in Panel A. TNA is the total of net assets under management and is reported in million USD. # Firms, # Countries, # Industries and # Supersectors are the number of distinct instances of the respective classification in the fund portfolio. We identify firms by their Worldscope Permanent Identifier and industries as well as supersectors by the respective ICB classification. We report the inverse of the normalized Herfindahl Index on the ICB industry classification. In Panel B, we report average portfolio weights across broad regions. The imposed regions follow the MSCI market classification: NAM = North America, EUR = Europe, JPN = Japan, APA = Asia Pacific excluding Japan, EM = Emerging Markets, FM = Frontier Market. Numbers in brackets denote p-values. Statistical significance at the 10%, 5%, and 1% level is indicated by \*, \*\*, and \*\*\*, respectively. The time period of analysis is from 2001 Q3 until 2014 Q1 (51 quarters).

	Crowding											
	1 (low)	2	3	4	5	6	7	8	9	10 (high)	10–1	
<b>Panel A: Fund characteristics</b>												
Centrality	28.59	65.52	110.17	164.08	238.13	316.47	389.95	460.75	534.89	649.18	620.59***	(0.000)
TNA	302	481	531	572	620	702	822	1,024	994	859	557.06***	(0.000)
# Firms	107	168	125	115	118	109	118	148	194	322	214.88***	(0.000)
# Countries	6	8	8	9	10	11	11	11	12	13	6.21***	(0.000)
# Industries	8	9	9	9	9	9	9	10	10	10	1.49***	(0.000)
# Supersector	15	15	15	16	15	15	16	17	17	18	3.41***	(0.000)
Inverse normalized HFI (industries)	14.86	21.69	22.12	16.61	58.61	35.36	24.62	26.29	27.47	31.00	16.15***	(0.000)
<b>Panel B: Weights for stock region</b>												
NAM	51.9	61.0	50.3	43.6	44.4	48.7	55.1	59.1	60.6	68.1	16.19***	(0.002)
EUR	22.6	15.1	11.5	17.1	35.2	42.0	36.9	33.4	32.7	26.3	3.71	(0.355)
APA	3.8	7.7	11.7	9.9	6.4	2.6	2.3	2.0	1.8	1.5	−2.33***	(0.000)
JPN	6.1	4.6	13.0	16.7	7.3	4.9	4.2	4.1	3.8	3.5	−2.63***	(0.000)
EM	13.0	9.3	11.9	11.4	5.9	1.3	1.1	0.9	0.7	0.5	−12.46***	(0.000)
FM	2.7	2.3	1.6	1.4	0.8	0.5	0.4	0.3	0.3	0.2	−2.48***	(0.000)

**Table 3:** Summary statistics of stock characteristics for decile portfolios sorted on crowding

This table provides stock summary statistics of fund portfolios conditional on crowding. In Panel A, we report the following variables for all stocks in each decile portfolio: Size is the log of primary issue market capitalization in billion USD, BTM is log of industry-adjusted book-to-market ratio; Momentum is the 9-month return preceeding quarterly raw returns; # Analysts is the number of analysts following the stock in the IBES database; Volatility is the volatility of monthly returns during the last 12 months; Price is in USD; ADR is an indicator variable taking 1 if the stock is an American Depositary Receipt and 0 otherwise; MSCI is an indicator variable taking 1 if the stock is part of the MSCI World Index and 0 otherwise; Anti-Director Index is a measure of shareholder protection based on Porta et al. (1998); Foreign Ownership is an indicator variable taking 0 if the fund's and stock issuer's domicile coincide and 1 otherwise. In Panel B, we report the following measures based on Sarkissian and Schill (2003) regarding the relation between the fund's and the stock issuer's domicile: Cultural Proximity is an indicator variable taking 1 if both countries share a common major spoken language or if they were part of the same colonial empire and 0 otherwise; Geographic Proximity is the distance between both countries' capitals in 1,000 kilometers; Economic Proximity is the percent of exports from the fund's country of domicile to the stock issuer's country of domicile. Numbers in brackets denote p-values. Statistical significance at the 10%, 5%, and 1% level is indicated by \*, \*\*, and \*\*\*, respectively.

	Crowding										
	1 (low)	2	3	4	5	6	7	8	9	10 (high)	10 - 1
<b>Panel A: all stocks</b>											
Size	3.22	6.67	14.93	25.98	39.43	49.23	55.89	65.04	75.72	85.28	82.06*** (0.001)
BTM (industry-adjusted)	0.08	-0.07	-0.18	-0.19	-0.19	-0.22	-0.24	-0.26	-0.26	-0.26	-0.34*** (0.001)
Momentum	0.24	0.25	0.22	0.20	0.15	0.15	0.15	0.14	0.13	0.12	-0.12*** (0.001)
# Analysts	10.08	12.79	16.68	19.38	23.44	25.73	26.47	26.80	27.81	28.58	18.50*** (0.001)
Dividend Yield	1.54	1.47	1.68	1.86	2.15	2.05	2.16	2.16	2.15	2.16	0.62*** (0.001)
Amihud Illiquidity	0.51	0.10	0.37	0.13	0.03	0.03	0.03	0.02	0.02	0.02	-0.49** (0.012)
Volatility	0.39	0.37	0.34	0.32	0.30	0.30	0.28	0.28	0.27	0.26	-0.12*** (0.001)
Turnover	0.16	0.18	0.17	0.15	0.13	0.14	0.14	0.14	0.14	0.13	-0.03*** (0.006)
Price	47.74	64.65	102.32	163.33	303.58	436.14	269.09	293.19	297.58	190.84	0.14*** (0.001)
ADR	0.02	0.03	0.04	0.05	0.05	0.04	0.03	0.03	0.02	0.02	0.00 (0.696)
MSCI	0.10	0.22	0.44	0.54	0.68	0.78	0.83	0.85	0.88	0.91	0.81*** (0.001)
English Legal Origin	0.79	0.82	0.76	0.73	0.74	0.76	0.79	0.82	0.83	0.86	0.07*** (0.001)
Anti-Director Index	3.41	3.31	3.44	3.55	3.53	3.42	3.32	3.27	3.22	3.19	-0.22*** (0.000)
Foreign Ownership	0.40	0.48	0.57	0.61	0.61	0.63	0.59	0.59	0.64	0.68	0.28*** (0.001)
<b>Panel B: foreign stocks</b>											
Cultural Proximity	0.34	0.33	0.32	0.30	0.32	0.33	0.34	0.34	0.31	0.27	-0.06*** (0.002)
Geographic Proximity	3.85	4.41	5.39	5.35	4.51	4.03	4.01	4.11	4.08	4.13	0.28** (0.045)
Economic Proximity	9.26	8.31	9.33	10.67	11.74	11.79	10.60	10.03	9.84	8.50	-0.76 (0.262)

**Table 4:** The performance of funds sorted on crowding

This table provides average monthly fund performance for portfolios conditional on crowding. At the end of each quarter, we sort funds into ten portfolios, based on crowding. Next, we track the equal-weighted performance of the portfolios during the next three months, when we rebalance. As fund performance measures, we use Net Alpha, Dollar Value Added, Gross Alpha, and DGTW returns. All performance measures are reported on a monthly basis and in percent, with the exception of Dollar Value Added which is reported in billion USD. We report time-series averages with t-statistics in parentheses based on Newey-West standard errors. Statistical significance at the 10%, 5%, and 1% level is indicated by \*, \*\*, and \*\*\*, respectively.

	Crowding										
	1 (low)	2	3	4	5	6	7	8	9	10 (high)	10 – 1
Net Alpha	0.102* (1.66)	0.042 (1.21)	0.002 (0.07)	-0.034 (-0.90)	-0.030 (-0.91)	-0.098*** (-3.43)	-0.100*** (-3.66)	-0.108*** (-4.61)	-0.116*** (-4.57)	-0.114*** (-4.54)	-0.215*** (-3.31)
Dollar Value Added	-0.009 (-0.02)	0.611** (2.43)	0.058 (0.21)	0.228 (0.68)	0.407 (1.18)	-0.921* (-1.76)	-0.037 (-0.08)	-0.216 (-0.44)	-0.687 (-1.20)	-1.855** (-2.19)	-1.846* (-1.83)
Gross Alpha	-0.065 (-1.50)	-0.063* (-1.72)	-0.071* (-1.89)	-0.109*** (-2.96)	-0.096*** (-2.87)	-0.115*** (-3.74)	-0.141*** (-4.83)	-0.160*** (-5.45)	-0.161*** (-5.82)	-0.156*** (-5.89)	-0.092*** (-3.08)
Gross DGTW	0.090 (1.44)	0.088 (1.21)	0.089 (1.32)	0.032 (0.55)	0.034 (0.73)	-0.018 (-0.38)	-0.020 (-0.55)	-0.034 (-1.09)	-0.057* (-1.94)	-0.046* (-1.92)	-0.136** (-2.49)

**Table 5:** Predictive regressions of fund performance on fund size and crowding

This table presents the results of predictive regressions of monthly fund returns on log of crowding and log fund size. Crowding is measured by either *crowd* or *PeerTNA*, which are defined according to equations 4 and 6, respectively. Fund sizes and PeerTNAs are inflated to millions of 2014 USD dollars using the value of all stocks in our sample and scaled by  $10^6$  in order to make coefficients easier to read. The estimator used in these regressions is defined in Section 4.2. In Panel A, the dependent variable is Net Alpha, in Panel B – Gross Alpha, and in Panel C – characteristics-adjusted returns (Gross DGTW). We report t-statistics in parentheses based on robust standard errors clustered on the fund and month level. \* denotes significance at the 10% level, \*\* at the 5% level, and \*\*\* at the 1% level.

	Model 1	Model 2	Model 3	Model 4	Model 5
<b>Panel A: Net Alpha</b>					
Ln( <i>crowd</i> <sup>w</sup> )	-0.0023*** (-3.31)			-0.0020*** (-2.68)	
Ln( <i>PeerSize</i> )		-0.0024*** (-2.60)			-0.0021** (-2.43)
Ln( <i>FundSize</i> )			-0.0010 (-1.35)	-0.0011 (-1.43)	-0.0009 (-1.17)
Number of Observations	450,387	450,387	450,387	450,387	450,387
<b>Panel B: Gross Alpha</b>					
Ln( <i>crowd</i> <sup>w</sup> )	-0.0014** (-2.54)			-0.0009** (-2.06)	
Ln( <i>PeerSize</i> )		-0.0005 (-1.00)			-0.0004 (-1.07)
Ln( <i>FundSize</i> )			-0.0012* (-1.80)	-0.0013* (-1.87)	-0.0012* (-1.73)
Number of Observations	454,671	454,671	454,671	454,671	454,671
<b>Panel C: Gross DGTW</b>					
Ln( <i>crowd</i> <sup>w</sup> )	-0.0022*** (-2.94)			-0.0019** (-2.10)	
Ln( <i>PeerSize</i> )		-0.0037*** (-4.35)			-0.0034*** (-4.19)
Ln( <i>FundSize</i> )			-0.0009 (-0.92)	-0.0010 (-1.05)	-0.0007 (-0.80)
Number of Observations	454,671	454,671	454,671	454,671	454,671

**Table 6:** Performance of funds in sequential sorts on size and crowding

This table provides average monthly fund performance for portfolios conditional on size and crowding. At the end of each quarter, we sort funds into ten portfolios, based on size. Next, we sort funds within each size decile into three portfolios based on crowding. Next, we track the equal-weighted net alpha of the portfolios during the next three months, when we rebalance. Net alpha is reported on a monthly basis and in percent. We report time-series averages with t-statistics in parentheses based on Newey-West standard errors. Statistical significance at the 10%, 5%, and 1% level is indicated by \*, \*\*, and \*\*\*, respectively.

Crowding	Fund Size										
	1 (low)	2	3	4	5	6	7	8	9	10 (high)	10 – 1
High	−0.194*** (−6.03)	−0.150*** (−6.04)	−0.147*** (−5.18)	−0.133*** (−5.07)	−0.131*** (−4.83)	−0.102*** (−4.02)	−0.100*** (−4.04)	−0.107*** (−3.93)	−0.073*** (−3.39)	−0.074*** (−3.51)	0.119*** (4.72)
Medium	−0.145*** (−4.31)	−0.062** (−2.00)	−0.087** (−2.42)	−0.074** (−2.21)	−0.079** (−2.17)	−0.062** (−2.04)	−0.070** (−2.03)	−0.062** (−2.19)	−0.021 (−0.71)	−0.020 (−0.86)	0.125*** (4.41)
Low	−0.068* (−1.71)	0.000 (0.00)	0.049 (1.18)	0.031 (0.70)	0.020 (0.44)	0.033 (0.68)	0.066* (1.68)	0.074* (1.89)	0.063** (2.05)	0.104*** (3.72)	0.171*** (4.06)
High – Low	−0.126** (−2.56)	−0.150*** (−3.50)	−0.196*** (−4.56)	−0.165*** (−3.40)	−0.151*** (−3.04)	−0.135** (−2.59)	−0.165*** (−3.73)	−0.181*** (−3.72)	−0.136*** (−4.06)	−0.178*** (−5.35)	

**Table 7:** Regressions of Fund Flow on PeerFlow

This table reports the results from regressions of quarterly fund performance on PeerFlow. At the end of each quarter, we compute quarterly fund performance using either net alpha, gross returns, or DGTW returns. All performance measures are expressed in a monthly basis and in percent. In Panel A, we use lagged PeerFlow as main explanatory variable, and in Panel B employ a 2SLS estimator with contemporaneous PeerFlow instrumented by lagged PeerFlow. In both Panels we include lagged control variables: an indicator variable TopCrowd taking the value of 1 if the fund is among the top 30% percent of crowding that quarter and 0 otherwise, log of crowding, log of fund size, fund flow, the most recently available fund expense ratio and turnover. Fund sizes and PeerFlows are inflated to millions of 2014 USD dollars using the value of all stocks in our sample and scaled by  $10^6$  in order to make coefficients easier to read. All specifications include fund and quarter fixed effects. We report t-statistics in parentheses based on robust standard errors clustered on the fund and quarter level. \* denotes significance at the 10% level, \*\* at the 5% level, and \*\*\* at the 1% level.

<b>Panel A:</b> Predictive relationship						
	NetAlpha <sub>t+1</sub>		GrossAlpha <sub>t+1</sub>		DGTW <sub>t+1</sub>	
PeerFlow <sub>t</sub>	17.2139 (0.39)	10.3426 (0.23)	-11.8694 (-0.31)	-19.0413 (-0.47)	-6.8134 (-0.12)	-22.9517 (-0.39)
PeerFlow <sub>t</sub> × TopCrowd <sub>t</sub>		45.9080 (1.63)		47.7870** (2.26)		106.4876*** (3.32)
TopCrowd <sub>t</sub>		0.0014 (1.65)		0.0014* (1.94)		0.0020* (1.97)
Ln(Crowd <sup>w</sup> ) <sub>t</sub>	-0.0068*** (-3.05)	-0.0071*** (-2.97)	-0.0042*** (-2.80)	-0.0046*** (-2.85)	-0.0072*** (-3.69)	-0.0076*** (-3.65)
Ln(FundSize) <sub>t</sub>	-0.0049*** (-7.65)	-0.0049*** (-7.64)	-0.0021*** (-4.33)	-0.0021*** (-4.29)	-0.0040*** (-5.31)	-0.0040*** (-5.22)
Flow <sub>t</sub>	-0.0030* (-1.97)	-0.0028* (-1.91)	-0.0026* (-1.98)	-0.0024* (-1.86)	-0.0034** (-2.07)	-0.0030* (-1.90)
ExpRatio <sub>t</sub>	-0.1088 (-1.43)	-0.1093 (-1.43)	0.0218 (0.39)	0.0214 (0.38)	-0.0052 (-0.12)	-0.0038 (-0.09)
Turn <sub>t</sub>	-0.0000 (-0.39)	-0.0000 (-0.37)	-0.0000 (-0.79)	-0.0000 (-0.68)	-0.0000 (-0.03)	0.0000 (0.01)
Observations	94,056	94,056	94,587	94,587	104,911	104,911
R2	0.11	0.11	0.12	0.12	0.16	0.16
Method	OLS	OLS	OLS	OLS	OLS	OLS
<b>Panel B:</b> Contemporaneous relationship						
	NetAlpha <sub>t+1</sub>		GrossAlpha <sub>t+1</sub>		DGTW <sub>t+1</sub>	
PeerFlow <sub>t+1</sub>	67.8172 (0.79)	60.7656 (0.70)	7.9489 (0.11)	-1.0499 (-0.01)	45.9536 (0.42)	27.3012 (0.24)
PeerFlow <sub>t+1</sub> × TopCrowd <sub>t</sub>		65.9361 (1.22)		84.1325* (1.86)		175.9613** (2.49)
TopCrowd <sub>t</sub>		0.0019* (1.85)		0.0024*** (2.78)		0.0043*** (3.49)
Ln(Crowd <sup>w</sup> ) <sub>t</sub>	-0.0058** (-2.67)	-0.0062*** (-2.69)	-0.0035** (-2.47)	-0.0040** (-2.66)	-0.0063*** (-3.34)	-0.0071*** (-3.51)
Ln(FundSize) <sub>t</sub>	-0.0053*** (-7.70)	-0.0053*** (-7.68)	-0.0023*** (-4.42)	-0.0023*** (-4.40)	-0.0043*** (-5.05)	-0.0043*** (-5.01)
Flow <sub>t</sub>	-0.0026* (-1.88)	-0.0025* (-1.86)	-0.0021* (-1.71)	-0.0019 (-1.63)	-0.0039** (-2.47)	-0.0036** (-2.38)
ExpRatio <sub>t</sub>	-0.1311* (-1.83)	-0.1313* (-1.83)	0.0016 (0.03)	0.0015 (0.03)	-0.0048 (-0.12)	-0.0032 (-0.08)
Turn <sub>t</sub>	-0.0000 (-0.40)	-0.0000 (-0.31)	-0.0000 (-1.23)	-0.0000 (-0.75)	-0.0000 (-0.39)	-0.0000 (-0.16)
Observations	97,611	97,611	35 97,908	97,908	108,193	108,193
R2	0.11	0.11	0.12	0.12	0.16	0.17
Method	IV	IV	IV	IV	IV	IV

**Table 8:** Predictive regressions of stock demand on characteristics, conditional on crowding

This table reports the results from predictive regressions of stock demand on stock characteristics, conditional on crowding. At the end of each quarter, we sort funds into ten portfolios, based on crowding. Next, using funds in each portfolio, we compute the demand for stocks they trade this quarter. For each portfolio and each quarter, we regress stock demand on lagged demand and lagged characteristics. Size is the log of primary issue market capitalization in billion USD, btm is log of industry-adjusted book-to-market ratio; Momentum is the 12-month return preceeding quarterly raw returns; Amihud illiquidity as in Amihud (2002), Analysts is the number of analysts following the stock in the IBES database; Volatility is the volatility of monthly returns during the last 12 months; Dividend Yield is the annual dividend yield; MSCI is an indicator variable taking 1 if the stock is part of the MSCI World Index and 0 otherwise. We report time-series averages of estimated coefficients with t-statistics in parentheses based on Newey-West standard errors. \* denotes significance at the 10% level, \*\* at the 5% level, and \*\*\* at the 1% level.

Dependent Variable: Demand <sub>t+1</sub>										
	Crowding									
	1 (low)	2	3	4	5	6	7	8	9	10 (high)
Demand <sub>t</sub>	0.500*** (18.72)	0.466*** (18.05)	0.439*** (21.59)	0.391*** (15.79)	0.421*** (21.73)	0.417*** (20.98)	0.361*** (11.63)	0.381*** (12.19)	0.381*** (11.86)	0.394*** (11.87)
Size <sub>t</sub>	0.007* (1.95)	0.011 (1.61)	0.027*** (4.04)	0.029*** (4.89)	0.009 (1.51)	0.015* (1.94)	0.031*** (4.43)	0.001 (0.09)	0.012 (1.10)	0.033*** (4.24)
Btm <sub>t</sub>	0.014*** (3.36)	0.010*** (3.29)	0.018*** (5.47)	0.007*** (3.46)	0.009*** (3.44)	0.010*** (3.66)	0.001 (0.61)	0.008** (2.43)	0.009*** (2.87)	0.000 (-0.11)
Momentum <sub>t</sub>	-0.006 (-1.26)	0.044*** (6.38)	0.045*** (3.68)	0.056*** (6.35)	0.043*** (3.52)	0.069*** (4.57)	0.099*** (8.98)	0.088*** (7.34)	0.098*** (8.60)	0.120*** (8.54)
Amihud Illiquidity <sub>t</sub>	-0.207 (-1.36)	-0.732* (-1.85)	-0.421 (-0.55)	-3.789** (-2.32)	-2.743** (-2.58)	-9.628*** (-2.73)	-9.084*** (-3.51)	-4.671*** (-2.72)	-8.916*** (-2.69)	-20.666*** (-3.63)
Volatility <sub>t</sub>	-0.064*** (-3.52)	-0.098*** (-3.14)	-0.093*** (-3.40)	-0.113*** (-3.04)	-0.204*** (-5.46)	-0.152*** (-4.72)	-0.112** (-2.60)	-0.124*** (-3.32)	-0.088** (-2.11)	-0.038 (-1.07)
Analysts <sub>t</sub>	-0.003*** (-4.53)	-0.002*** (-3.61)	-0.002*** (-6.14)	-0.002*** (-4.22)	-0.002*** (-3.82)	-0.001*** (-5.87)	-0.001** (-2.41)	-0.001*** (-5.03)	-0.001** (-2.06)	-0.002*** (-4.79)
Dividend Yield <sub>t</sub>	0.003*** (2.69)	-0.002** (-2.31)	0.002* (1.68)	0.002 (1.01)	0.003*** (2.68)	0.000 (-0.14)	-0.002* (-1.69)	-0.007*** (-3.44)	-0.012*** (-5.16)	-0.007*** (-3.93)
MSCI <sub>t</sub>	-0.077*** (-5.99)	-0.040*** (-3.08)	-0.055*** (-4.68)	-0.028** (-2.15)	-0.007 (-0.64)	0.008 (0.79)	-0.002 (-0.12)	0.009 (0.77)	0.008 (0.48)	0.027* (1.85)
Observations	408,398	352,701	319,817	276,809	253,365	219,969	203,281	230,517	246,310	234,167
R2	0.28	0.25	0.22	0.19	0.22	0.21	0.18	0.20	0.22	0.23

**Table 9: Liquidity**

At the end of each quarter, we sort funds into ten portfolios, based on crowding. Next, we track the equal-weighted performance of the excess returns of portfolios during the next three months, when we rebalance. We obtain a time-series of portfolio returns. In Panel A, we use a Fama-French model to obtain alpha. In Panel B, we add the liquidity factor of Pastor and Stambaugh (2003). We report estimated alphas and betas, with t-statistics in parentheses. \* denotes significance at the 10% level, \*\* at the 5% level, and \*\*\* at the 1% level.

<b>Panel A: 3 factor loadings</b>											
	1 (low)	2	3	4	5	Crowding		8	9	10 (high)	10 – 1
						6	7				
Alpha	0.098 (1.01)	−0.003 (−0.05)	0.009 (0.08)	0.001 (0.01)	−0.004 (−0.03)	−0.074 (−0.88)	−0.062 (−0.81)	−0.098 (−1.51)	−0.129** (−2.50)	−0.116*** (−3.51)	−0.214*** (−2.79)
MKT beta	0.983*** (25.67)	1.028*** (36.85)	1.050*** (27.54)	1.068*** (23.02)	1.024*** (28.50)	1.016*** (55.64)	0.983*** (42.49)	0.972*** (41.04)	0.962*** (49.35)	0.916*** (87.05)	−0.067** (−2.22)
SMB beta	0.636*** (15.75)	0.583*** (16.28)	0.407*** (10.87)	0.246*** (4.39)	0.161*** (3.58)	0.032 (0.93)	−0.035 (−1.01)	−0.065** (−2.22)	−0.090*** (−3.96)	−0.119*** (−8.80)	−0.756*** (−20.60)
HML beta	0.128** (2.18)	0.039 (1.08)	−0.005 (−0.10)	−0.105* (−1.72)	−0.145** (−2.52)	−0.087** (−2.38)	−0.034 (−0.97)	−0.013 (−0.40)	−0.035 (−1.59)	−0.025* (−1.78)	−0.154*** (−2.91)
<b>Panel B: Liquidity factor loadings</b>											
	1 (low)	2	3	4	5	Crowding		8	9	10 (high)	10 – 1
						6	7				
Alpha	0.040 (0.43)	−0.048 (−0.67)	−0.057 (−0.59)	−0.076 (−0.60)	−0.086 (−0.80)	−0.118 (−1.48)	−0.101 (−1.40)	−0.126** (−1.96)	−0.139*** (−2.62)	−0.126*** (−3.70)	−0.166** (−2.27)
MKT beta	0.963*** (28.05)	1.012*** (38.09)	1.027*** (31.17)	1.042*** (24.95)	0.995*** (33.09)	1.000*** (63.62)	0.969*** (46.73)	0.962*** (43.98)	0.959*** (51.42)	0.913*** (95.23)	−0.051* (−1.82)
SMB beta	0.616*** (14.68)	0.567*** (15.70)	0.384*** (10.01)	0.219*** (3.91)	0.132*** (3.01)	0.016 (0.49)	−0.048 (−1.42)	−0.075** (−2.51)	−0.094*** (−3.95)	−0.123*** (−8.73)	−0.739*** (−19.24)
HML beta	0.166*** (3.34)	0.068** (2.09)	0.038 (1.00)	−0.055 (−1.12)	−0.093** (−2.06)	−0.058* (−1.89)	−0.009 (−0.30)	0.006 (0.20)	−0.028 (−1.31)	−0.019 (−1.39)	−0.184*** (−4.00)
Liquidity beta	0.094*** (4.40)	0.073*** (3.39)	0.106*** (5.35)	0.124*** (4.64)	0.132*** (5.11)	0.072*** (3.65)	0.063*** (3.67)	0.046*** (2.75)	0.017 (1.07)	0.017* (1.67)	−0.077*** (−4.86)



**Table 10:** Performance of funds in sorts on crowding and momentum trading

This table provides average monthly fund performance for portfolios conditional on crowding and momentum trading. At the end of each quarter, we sort funds into ten portfolios, based on crowding. Next, we independently group time periods into three buckets, based on the estimated coefficient on momentum in Equation (10). In Panel A, we report the equal-weighted net alpha of the portfolios during the next three months. In Panel B, we report the equal-weighted net alpha of the portfolios during the next twelve months. Net alpha is reported on a quarterly (yearly) basis and in percent in Panel A (B). We report time-series averages with t-statistics in parentheses based on Newey-West standard errors. Statistical significance at the 10%, 5%, and 1% level is indicated by \*, \*\*, and \*\*\*, respectively.

<b>Panel A: 3-months alpha</b>										
Demand for Momentum	Crowding									
	1 (low)	2	3	4	5	6	7	8	9	10 (high)
1 (low)	0.474*** (3.27)	-0.119 (-0.72)	0.035 (0.21)	-0.113 (-1.33)	0.091 (0.98)	-0.167 (-1.73)	-0.209 (-1.73)	-0.250*** (-3.10)	-0.151* (-1.90)	-0.291*** (-3.00)
2	0.259 (0.90)	0.373 (1.24)	0.379*** (3.05)	0.289* (1.93)	-0.089 (-0.60)	-0.228 (-1.54)	-0.228** (-2.38)	-0.249** (-2.46)	-0.309*** (-5.95)	-0.386*** (-10.39)
3 (high)	0.438* (1.91)	0.405*** (3.10)	-0.163 (-1.54)	-0.163 (-0.80)	0.059 (0.64)	-0.247*** (-3.35)	-0.181 (-1.59)	-0.299*** (-3.46)	-0.418** (-2.84)	-0.247* (-1.83)
<b>Panel B: 12-months alpha</b>										
Demand for Momentum	Crowding									
	1 (low)	2	3	4	5	6	7	8	9	10 (high)
1 (low)	2.072*** (6.62)	0.279 (0.80)	0.594** (2.33)	-0.085 (-0.16)	-0.374 (-1.10)	-0.692 (-1.37)	-1.186*** (-5.44)	-1.182*** (-7.37)	-1.471*** (-13.07)	-1.335*** (-5.83)
2	0.680 (0.90)	0.510 (1.27)	0.762** (2.30)	0.872* (1.96)	-0.597 (-1.74)	-0.904** (-2.39)	-0.786*** (-4.24)	-0.857*** (-9.43)	-1.184*** (-8.80)	-1.226*** (-5.07)
3 (high)	1.944 (1.68)	0.671** (2.50)	-0.272 (-1.01)	-0.930** (-2.20)	0.267 (0.66)	-0.773*** (-4.07)	-0.695*** (-3.38)	-0.800*** (-5.18)	-1.068** (-2.83)	-1.046*** (-3.55)

**Table 11:** Performance of funds in sequential sorts on active share and crowding

This table provides average monthly fund performance for portfolios conditional on active share and crowding. At the end of each quarter, we sort funds into ten portfolios, based on active share. Active share follows the definition of Cremers and Petajisto (2009) where we use the MSCI World Index as the benchmark. Next, we sort funds within each size decile into three portfolios based on crowding. Next, we track the equal-weighted net alpha of the portfolios during the next three months, when we rebalance. Net alpha is reported on a monthly basis and in percent. We report time-series averages with t-statistics in parentheses based on Newey-West standard errors. Statistical significance at the 10%, 5%, and 1% level is indicated by \*, \*\*, and \*\*\*, respectively.

Active Share	Crowding										
	1 (low)	2	3	4	5	6	7	8	9	10 (high)	10 – 1
High	−0.102*** (−4.01)	−0.157*** (−5.48)	−0.177*** (−5.36)	−0.165*** (−3.83)	−0.178*** (−3.95)	−0.078* (−1.85)	0.027 (0.57)	−0.018 (−0.26)	0.018 (0.23)	0.074 (1.54)	0.176*** (3.07)
Medium	−0.087*** (−3.54)	−0.112*** (−5.26)	−0.128*** (−4.76)	−0.134*** (−4.36)	−0.116*** (−3.41)	−0.054* (−1.76)	−0.065** (−2.19)	−0.009 (−0.26)	0.089** (2.02)	0.003 (0.07)	0.090* (1.90)
Low	−0.021 (−0.62)	−0.036 (−1.46)	−0.059** (−2.50)	−0.088*** (−3.76)	−0.070*** (−2.76)	−0.075** (−2.13)	−0.055 (−1.56)	0.056 (1.10)	0.119* (1.69)	0.181* (1.69)	0.202* (1.89)
High – Low	−0.081*** (−3.01)	−0.121*** (−4.52)	−0.118*** (−4.40)	−0.077* (−1.92)	−0.108** (−2.53)	−0.003 (−0.08)	0.083* (1.78)	−0.074 (−0.91)	−0.101 (−0.93)	−0.107 (−0.97)	

**Table 12:** Portfolio weights in broad regions conditional on crowding

This table provides the average portfolio weights for three fund portfolios by the stock region. In each quarter, funds are sorted into ventile portfolios, for each of which we compute cross-section averages of region portfolio weights. We report the quarterly mean weights for the top and bottom portfolios as well as the universe of funds. Next, we compute the average difference between the region weights of the MSCI World Index and the bottom and top crowding ventiles as well as the universal fund portfolio. The imposed regions follow the MSCI market classification: NAM = North America, EUR = Europe, JPN = Japan, APA = Asia Pacific excluding Japan, EM = Emerging Markets, FM = Frontier Market. Numbers in parantheses are Newey-West corrected p-values. Statistical significance at the 10%, 5%, and 1% level is indicated by \*, \*\*, and \*\*\*, respectively.

Stock issuer	Portfolio weights				Portfolio weights minus MSCI World Index weights					
	Universe	Bottom	Top	MSCI	Universe		Bottom		Top	
NAM	69.1	47.0	80.0	55.5	13.61***	(0.00)	−8.54***	(0.00)	24.50***	(0.00)
EUR	18.5	22.6	15.1	30.1	−11.61***	(0.00)	−7.56***	(0.00)	−15.02***	(0.00)
APA	3.5	4.5	1.4	4.4	−0.83***	(0.00)	0.15	(0.46)	−2.92***	(0.00)
JPN	4.1	8.7	2.4	9.8	−5.78***	(0.00)	−1.18**	(0.03)	−7.39***	(0.00)
EM	3.9	15.1	0.8	0.0	3.92***	(0.00)	15.08***	(0.00)	0.78***	(0.00)
FM	0.8	2.2	0.2	0.1	0.69***	(0.00)	2.04***	(0.00)	0.04	(0.11)

**Table 13:** Fama-MacBeth regressions of net alpha on lagged crowding

This table reports results from Fama-MacBeth regressions of fund net alpha on crowding. The dependent variable is net alpha averaged over the three months following any portfolio holding report (time  $t + 1$ ).  $crowd$  is crowding based on portfolio holdings overlap and  $crowd^e$  is crowding measured as the eigenvector centrality following the definition in equations (4) and (11), respectively.  $QuarterlyFlow_t$  is the fund flow from quarter  $t - 1$  to  $t$  computed as  $(TNA_t - TNA_{t-1}(1 + R_t))/TNA_{t-1}$ .  $TNA_t$  is in millions and  $\log(TNA_t)$  is the log thereof.  $AverageNetAlpha_t$  is Net Alpha averaged over the three months preceeding any portfolio report. Standard errors are Newey-West corrected. Numbers in parantheses are t-statistics. Statistical significance at the 10%, 5%, and 1% level is indicated by \*, \*\*, and \*\*\*, respectively.

	Model 1	Model 2	Model 3	Model 4
Constant	0.040 (1.18)	0.023 (0.68)	-0.006 (-0.14)	-0.022 (-0.58)
$crowd_t$	-0.000*** (-4.34)		-0.000*** (-3.30)	
$crowd_t^e$		-0.002*** (-3.68)		-0.002*** (-3.07)
$QuarterlyFlow_t$			-0.000 (-0.03)	-0.000 (-0.00)
$\log(TNA)_t$			0.007* (1.75)	0.007* (1.91)
$NetAlpha_t$			0.051* (1.91)	0.052** (2.01)
Number of observations	150,202	150,202	150,202	150,202
Method	FamaMacBeth	FamaMacBeth	FamaMacBeth	FamaMacBeth

**Table 14:** Fama-Macbeth regressions of net alpha on lagged crowding for U.S. domiciled funds

This table reports results from Fama-MacBeth regressions of fund net alpha on crowding for U.S. domiciled funds. The dependent variable is net alpha averaged over the three months following any portfolio holding report (time  $t + 1$ ). *crowd* is crowding based on portfolio holdings overlap and *crowd<sup>e</sup>* is crowding measured as the eigenvector centrality following the definition in equations (4) and (11), respectively.  $TNA_t$  is in millions and  $\log(TNA_t)$  is the log thereof.  $AverageNetAlpha_t$  is Net Alpha averaged over the three months preceeding any portfolio report. NPeers (number of fund peers) and TSIM (total similarity of fund peers) are competition measures from Hoberg et al. (2017). Standard errors are Newey-West corrected. Numbers in parantheses are t-statistics. Statistical significance at the 10%, 5%, and 1% level is indicated by \*, \*\*, and \*\*\*, respectively.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Constant	0.303* (0.05)	0.256** (0.03)	0.237*** (0.00)	0.204*** (0.00)	0.228*** (0.00)	0.201** (0.01)	0.228*** (0.00)	0.201** (0.01)
$crowd_t$	-0.001*** (0.00)		-0.001*** (0.00)		-0.001*** (0.00)		-0.001*** (0.00)	
$crowd_t^e$		-0.005*** (0.00)		-0.005*** (0.00)		-0.005*** (0.00)		-0.005*** (0.00)
$\log(\text{fund TNA})_t$			0.006 (0.22)	0.005 (0.36)	0.007 (0.25)	0.006 (0.32)	0.007 (0.25)	0.006 (0.33)
$QuarterlyFlow_t$			0.001 (0.29)	0.001 (0.28)	0.001 (0.27)	0.001 (0.27)	0.001 (0.27)	0.001 (0.28)
$QuarterlyReturn_{t-1}$			0.019 (0.35)	0.019 (0.35)	0.019 (0.38)	0.019 (0.37)	0.019 (0.35)	0.019 (0.36)
NPeers					0.000 (0.93)	-0.000 (0.79)		
TSIM							-0.000 (0.98)	-0.000 (0.72)
Number of observations	119,750	119,750	112,288	112,288	112,288	112,288	112,288	112,288

# Appendices

## Appendix A Data screens

The data cleaning of the Factset and Datastream/Worldscope databases closely follows Dyakov et al. (2019). In Factset, we drop any fund report in which a single security constitutes more than 25% of the total assets of the fund. Next, we exclude portfolio reports with apparent data errors where the total net asset value appears to bounce back close to its original value after a significant positive or negative spike.<sup>5</sup> In some cases, there exist more than one report per quarter or the report does not refer to end-of-quarter positions (i.e., February rather than March). In such cases, we always choose the portfolio snapshot closest to the end of the quarter and use the reported holdings as if they were reported at the end of the quarter. Not all of the funds in our sample (roughly 20%) have portfolio holdings available on a quarterly basis. For those funds, we carry-on the last quarter's report to the current one in order to obtain an unbroken quarterly sequence of portfolio positions.

We next follow the data-cleaning procedures prescribed in Ince and Porter (2006), Schmidt et al. (2011), and Dyakov et al. (2019). Specifically, we exclude a) stock issues with more than 20% difference in market capitalization between Datastream and Factset, b) stocks where a single fund is reported to own more than 25% of the shares, and c) stocks with some key information missing in either Factset or Datastream. A few funds increase their holdings by a factor of e.g. 100 only to decrease their holdings by a similar factor in the next reporting period. Such changes are apparent data errors and we exclude them using the same screen for individual holdings as the one used for large reversals of total reported assets mentioned above. A detailed overview on the stock-level country selection, the merging of Factset with Datastream, the cleaning of stock information from Datastream, and the construction of characteristic-adjusted returns for international stocks is available in the Online Appendix of Dyakov et al. (2019). We report extended descriptive statistics of for our sample, based on country of domicile in Table

---

<sup>5</sup>Specifically, we drop fund portfolio reports across two periods where reported total net assets (TNA) increase/decrease by a factor of more than 9 (quarter  $q - 1$  vs. quarter  $q$ ) and which is subsequently reversed by a factor of at least 4.5 in the opposite direction (quarter  $q$  vs. quarter  $q + 1$ ) while the increase/decrease across both periods does not exceed 4.5 in the original direction (quarter  $q + 1$  vs. quarter  $q - 1$ ).

A1.

## **Appendix B The alternative investment opportunity set**

In this Section, we describe the choice of traded funds as an alternative investment opportunity set. Following Berk and van Binsbergen (2015), we use Vanguard funds due to their popularity among investors and the low trading costs associated with trading them. We use the following selection criteria. First, we select only equity funds and drop Morningstar Global Categories that span specific sectors of the stock market, such as technology and health care. Next, within each Global Category we select the oldest fund(s), offered in USD, that span all stocks in the category. We drop some country specific funds that are not offered in USD and whose coverage is already spanned by other funds. The final selection of funds results in 7 domestic U.S. funds and 6 international funds. The NAM set of funds consists of the 7 U.S. funds. For all other regions, we use the three Global Equity index funds. For the EUR region, we add the European Equity index fund. For the APA region, we add the Asia-Pacific Equity fund. Similarly, for the EME region, we add the Emerging Markets equity fund. For the GLO region, we select all international funds (except for the Emerging Markets equity fund), and add the three main U.S. funds that track the S&P 500, small stocks, and value stocks.

Table A2 presents the full list of passive funds. Our choice of selecting separate index funds for broad geographical regions is based on (Fama and French, 2012) and (Griffin, 2002), who show that global markets are not integrated and risk-premia are driven by local rather than global forces. We do not include funds with a distinctive momentum focus, as they are not available to investors during the span of our study. For instance, Blackrock offers Europe-specific momentum ETFs only since the beginning of 2015.

The benchmark loadings in (1) are estimated by regressing the fund's net excess returns upon the excess net returns of the relevant benchmarks over the entire sample period that the fund is active. When we compute gross alpha, we use the gross return of the benchmarks, defined as net returns plus one twelfth of the reported net annual expense ratio. Because one of the two global funds is not available throughout our sample period, we estimate betas by using

an augmented basis of the factors where the factor returns are orthogonalized with respect to all other variables and missing returns are replaced with the mean of the orthogonalized factor. Alphas are then estimated by using the estimated betas and the augmented basis where we replace missing returns with zero. See the Appendix in Berk and van Binsbergen (2015) for more details on the procedure. The augmented basis is computed separately for benchmarks in each of the geographical regions.

## Appendix C Crowding based on eigenvalue centrality

We acknowledge that the *crowding score* of a fund might not just depend on the direct neighbors but also on the neighbors' neighbors. We therefore introduce a new measure of crowding, based on the eigenvalue centrality of funds in a global network of overlapping holdings. The network is of size  $n \times n$ , where  $n$  is the total number of funds. Each element of the network represents the portfolio holdings overlap between two funds, as in equation (3). We therefore re-estimate our results using the fund's *eigenvector centrality* as a measure for crowding. This approach takes more than just the first degree neighbors into account when ranking funds: The centrality of fund  $i$  is a function of its neighbors' centralities. Thus, each fund's *crowding score* is recursively determined using the score of each other fund in the portfolio overlap network. The eigenvector centrality of fund  $i$  is given by

$$crowd_i^e = \frac{1}{\lambda} \sum_{j=1}^n A_{j,i} crowd_j^e \quad (11)$$

where  $A_{j,i}$  is the edge from  $j$  to  $i$  and  $\lambda$  is a constant. This means a fund receives a higher *crowding score* if it is connected to other funds with a high *crowding score*. A fund with a large number of unimportant neighbors might have a smaller *crowding score* than one with a smaller number of more important neighbors.<sup>6</sup> The solution to the recursive problem (11) can

---

<sup>6</sup>Eigenvector centrality is also often used to determine influence in social networks. The PageRank algorithm from the Google Search Engine is also a popular example for *eigenvector centrality*.



be rewritten to a eigen decomposition problem

$$\lambda \mathbf{C}_e^T = \mathbf{A}^T \mathbf{C}_e^T \quad (12)$$

where  $\mathbf{C}_e = (ce_1, ce_2, \dots, ce_n)$  is the left eigenvector of adjacency matrix  $\mathbf{A}^T$  and  $\lambda$  the corresponding eigenvalue.<sup>7</sup> To obtain positive centrality measures, we choose the largest eigenvalue for calculating (11).

Results using eigenvalue centrality are similar to our baseline findings. In Table A3, we sort funds in ten deciles and examine subsequent performance. Using net and gross alpha, dollar value added, and DGTW returns, we find that performance is decreasing with crowding. Results are qualitatively similar to the ones in Table 4.

## Appendix D The determinants of crowding and its persistence

Table A4 presents the results of a panel regression of crowding on a number of characteristics. Crowding increases in portfolio size and the measures of diversification. The relationship between centrality and portfolio size is convex. However, for a given increase in portfolio size, the effect becomes smaller as the number of stocks increases. Increasing the share of investment in the largest markets, North America and Europe, increases the crowding score while shifting towards Asia Pacific excluding Japan decreases crowding. For robustness, we split the sample of funds into those domiciled in the USA (Model 3) and those domiciled elsewhere (Model 4) and perform the same analysis. The results are consistent but the effect on crowding of shifting into regions in which funds are not domiciled is larger than the average effect for the entire sample. The results for the crowding score based on the eigenvector centrality (Models 5-10) are qualitatively similar.

In Figure 1, we examine the persistence of crowding. At the end of each quarter, we sort

---

<sup>7</sup>Every finite graph, like the network graph above, can be represented by an adjacency matrix. The entries of this matrix are the weights of the edges of the graph. The eigenvalue decomposition approach has been applied since at least Landau (1895), who used network centrality for ranking of chess players: the relative importance of each chess player was determined not only by the count of the player's wins, but also on the relative importance of each of her opponents, and the opponents' opponents, and so on. Similarly, the centrality of each mutual fund in our network is recursively determined by its portfolio holdings similarity with competitors, and the links between the fund's competitors with their competitors, and so on.

funds into deciles based on their crowding score and compute a matrix of transition rates for the following quarter. The transition matrix reveals that persistence is large with more than 69% of funds staying in the same decile across all deciles. Moreover, persistence is stronger for deciles below the median compared to those above the median. There seems to be no obvious systemic relation between crowding decile and dropping out of the sample.

## **Appendix E Robustness of fund performance, conditional on crowding**

In Table 4, we find a negative relationship between crowding and subsequent performance. When we group funds in the separate portfolios, we use equal-weighting. Alternatively, we can value-weight funds. Results are reported in Panel A of Table A5. In addition, we can risk-adjust returns using traditional factor regressions instead of the passive index fund approach of Berk and van Binsbergen (2015). Results are reported in Panel B. In all sorts, we find the same repeating patterns: performance is monotonically decreasing in alpha, and performance becomes negative among funds operating in the most crowded deciles.

In Table 6, we double sort funds on size and crowding and find that within each size decile, net alpha is decreasing with crowding. We examine the robustness of this findings, and employ alternative performance measures. Results are reported in Panel A of Table A6. In addition, in Panel B we sort on eigenvalue centrality instead of crowding. In both panels, we show the performance of the spread portfolios of the top minus bottom tercile of crowding within each decile. In both Panels, we find predominantly negative values. Thus, the findings in Table 6 are robust to alternative performance measures and the alternative measure of crowding (eigenvalue centrality).

## **Appendix F Details on the diseconomies of scale regressions**

In equation (5), we related fund performance to lagged crowding and fund size. A standard fixed effects estimator requires the regressors to be strictly exogenous. As we argue in the main

body of the paper, this is not always the case. To address this problem, we follow Pastor et al. (2015) and Zhu (2018) and first eliminate the fixed effects  $\alpha_f$  by forward-demeaning equation (5). The forward-demeaned version of a variable  $x$  is defined as

$$\bar{x}_{it} = x_{it} - \frac{1}{T_i - t + 1} \sum_{s=t}^{T_i} x_{is}, \quad (13)$$

where  $T_i$  denotes the number of time periods for which fund  $i$  is observed. We then estimate the coefficients in equation (5) using a two-stage least squares (2SLS) approach, employing instruments that are plausibly uncorrelated with the forward-demeaned error term. The literature propose two potential instruments for the forward-demeaned size. The first one, suggested by Pastor et al. (2015), is the backward-demeaned version of fund size, where the backward-demeaned version of a variable  $i$  is defined as

$$\underline{x}_{i,t-1} = x_{i,t-1} - \frac{1}{t-1} \sum_{s=1}^{t-1} x_{i,s-1}. \quad (14)$$

In the first stage of the 2SLS, Pastor et al. (2015) propose a reduced form for the endogenous regressor, the fitted values of which are substituted into the forward-demeaned version of (5) without an intercept. Zhu (2018), however, argues that an intercept term should be included in the reduced forms, and we follow her recommendation. She further proposes an alternative instrument for the forward-demeaned size, namely lagged fund size  $q_{i,t-1}$  because it is correlated with the forward-demeaned lagged fund size and it is plausibly uncorrelated with the forward-demeaned error term. She finds it is a stronger instrument as it improves the fit of the first-stage regressions. Following Dyakov et al. (2019), we combine both instruments into one estimator, which should lead to more efficient inference. In order to minimize the impact of estimation error on our findings, we drop funds with less than 4 years of data.

Thus, in the first stage we regress forward-demeaned fund size upon its backward-demeaned version as well as forward-demeaned crowding. In the second stage, we regress forward-demeaned alpha on the fitted values from the first stage regression as well as forward-demeaned crowding. Both size and crowding are expressed in logs, but for simplicity of notation we drop log in the

equations below.

$$\bar{q}_{i,t-1} = \psi + \rho_1 \overline{crowd}_{it-1}^w + \rho_2 \underline{q}_{i,t-1} + \rho_3 q_{i,t-1} + v_{i,t}, \quad (15)$$

$$\bar{r}_{i,t} = \beta_1 \overline{crowd}_{i,t-1}^w + \beta_2 \bar{q}_{i,t-1}^* + \vartheta_{i,t}. \quad (16)$$

As results are consistent across specifications where we include either of the instruments in equation (15), we only report results using both instrumental variables.

## Appendix G Directional Crowding Measures

Our main portfolio overlap measure assumes a symmetric relation between any two funds. An asymmetric measure of the relationship between funds relaxes this assumption. Consider a value-investing fund A that competes with fund B, which specializes in both value stocks and small stocks. Because A faces competition in its sole investment style, fund A can become be more exposed to fund B than the more diversified fund B is exposed to A. We introduce asymmetry to our benchmark portfolio overlap measure by scaling the connection from fund  $i$  to fund  $j$  by the number of overlapping stocks relative to the total number of stocks in the portfolio of fund  $j$ . Then, the weight of the edge from fund  $i$  to  $j$ ,  $i \rightarrow j$  is a measure of  $j$ 's exposure to  $i$ . The edge is computed as

$$\tilde{e}_{ij} = \frac{|P_i \cap P_j|}{|P_j|} \sum_{k \in P_j} \min(\omega_i^k, \omega_j^k), \quad (17)$$

where  $P_i$  and  $P_j$  is the set of stocks held by fund  $i$  and  $j$ , respectively. Note that if  $P_i = P_j$  or  $P_i \cap P_j = \emptyset$ , then the overlap measure is identical to the benchmark in equation (3). In all other cases, the edge is directional and an overlap among a relatively higher (lower) number of stocks receives a higher (lower) weight.

We compute our crowding scores using this alternative definition of the network edges and replicate our main analysis. Table A7 shows that the alternative measure generates qualitatively similar results. Panel A shows that funds in the top decile of crowding, calculated as weighted

degree centrality, generate -0.078% ( $p=0.005$ ) in subsequent DGTW-adjusted monthly returns. The spread in performance between the top and bottom decile amounts to -0.173% ( $p=0.005$ ). Results using eigenvalue centrality are similar (Panel B).

**Table A1:** Descriptive statistics, conditional on domicile

This table provides summary statistics of fund stock holdings by funds' domicile. We report the number of unique funds, mean total net assets under management (TNA, in million USD), mean number of stocks in portfolios, mean net alpha (in %) and the mean portfolio weights for each stock region (in %). We first compute cross-section and subsequently quarterly means. The imposed regions follow the MSCI market classification: NAM = North America, EUR = Europe, APA = Asia Pacific excluding Japan, JPN = Japan, EM = Emerging Markets, FM = Frontier Market.

Fund country	Funds count	TNA mean	Holdings mean	Stock region					
				NAM	EUR	APA	JPN	EM	FM
North America									
	Funds count	TNA mean	Holdings mean	NAM	EUR	APA	JPN	EM	FM
Canada	1,461	319	125	72.6	14.2	4.2	5.5	2.9	0.6
United States	5,026	1,311	180	79.3	10.2	3.0	3.7	2.9	1.0
Europe									
	Funds count	TNA mean	Holdings mean	NAM	EUR	APA	JPN	EM	FM
Austria	192	80	103	41.5	35.7	5.4	8.6	8.0	0.8
Belgium	285	120	149	29.8	54.3	3.5	6.3	5.4	0.6
Switzerland	291	198	192	32.8	43.9	6.2	12.1	4.6	0.4
Germany	537	309	94	28.0	58.9	2.7	8.5	1.6	0.4
Denmark	187	137	139	36.8	30.4	7.7	12.5	11.5	1.2
Spain	459	77	83	28.5	62.6	0.8	6.3	1.6	0.2
Finland	125	153	88	31.2	48.5	4.9	6.5	7.9	1.0
France	860	275	113	16.6	70.7	2.7	7.0	2.7	0.3
United Kingdom	1,497	413	122	24.9	53.7	7.0	8.5	4.9	0.9
Greece	15	39	97	43.8	48.9	1.6	1.6	3.2	0.8
Ireland	842	281	186	38.6	29.2	9.5	13.3	8.2	1.2
Italy	286	226	113	35.2	43.4	5.7	10.0	5.0	0.6
Luxembourg	3,666	291	151	37.3	35.3	7.4	10.8	8.0	1.2
Netherlands	182	596	120	36.7	39.6	8.3	10.1	4.5	0.8
Norway	101	306	155	31.8	52.4	4.0	8.4	2.1	1.4
Portugal	46	53	91	38.2	47.6	2.6	3.7	5.0	2.8
Sweden	272	325	164	29.8	51.4	4.8	7.6	5.6	0.8
Asia Pacific excluding Japan									
	Funds count	TNA mean	Holdings mean	NAM	EUR	APA	JPN	EM	FM
Australia	22	494	78	33.9	15.8	31.7	4.0	14.1	0.4
Hong Kong	68	184	111	15.0	10.2	32.0	22.6	20.1	0.3
Singapore	48	71	80	18.2	11.0	26.9	20.0	23.4	0.5

*Continues on next page...*

Continued Table...

Fund country	Funds count	TNA mean	Holdings mean	Stock region					
				NAM	EUR	APA	JPN	EM	FM
Japan									
	Funds count	TNA mean	Holdings mean	NAM	EUR	APA	JPN	EM	FM
Japan	84	553	128	20.4	16.4	4.7	52.7	5.3	0.5
Emerging Markets									
	Funds count	TNA mean	Holdings mean	NAM	EUR	APA	JPN	EM	FM
Brazil	5	259	56	0.0	2.0	0.0	0.0	98.0	0.0
Chile	12	64	68	44.5	0.5	1.8	0.0	52.4	0.8
China	57	1,016	110	7.1	1.1	13.4	0.3	78.1	0.1
Czech Republic	9	47	74	46.1	40.0	0.8	1.5	10.4	1.2
Hungary	5	73	85	32.6	45.8	5.3	4.8	10.3	1.2
Indonesia	8	292	66	0.4	0.0	0.2	0.0	99.4	0.0
India	134	189	58	0.7	0.3	1.9	0.0	97.0	0.1
Malaysia	36	42	70	5.6	5.1	27.0	3.9	58.1	0.2
Poland	109	160	76	5.1	6.8	0.1	0.1	86.7	1.2
Thailand	1	21	126	52.1	33.6	3.7	9.8	0.0	0.8
Taiwan	143	66	68	11.0	9.2	15.0	5.2	57.9	1.6
Frontier Markets									
	Funds count	TNA mean	Holdings mean	NAM	EUR	APA	JPN	EM	FM
Bermuda	4	549	55	6.4	9.9	20.5	4.3	58.2	0.7
Cayman Islands	52	188	179	21.0	11.3	29.7	8.8	27.7	1.5
Estonia	2	42	53	5.2	10.0	0.0	0.0	80.8	4.0
Gibraltar	1	29	97	14.4	29.1	20.0	0.1	32.0	4.5
Israel	55	33	73	14.7	4.2	1.6	0.7	2.9	76.0
Liechtenstein	60	115	113	36.5	32.7	7.2	19.3	3.9	0.4
Lithuania	1	28	53	2.2	69.1	0.0	0.0	13.4	15.2
Mauritius	4	285	54	2.6	0.2	1.7	0.0	95.4	0.0
Pakistan	3	350	134	0.0	0.0	0.0	0.0	0.0	100.0
Slovakia	3	43	124	52.3	42.9	0.8	3.9	0.0	0.2
Slovenia	31	44	73	21.6	36.1	2.7	1.7	5.1	32.8
British Virgin Islands	3	55	82	23.8	71.7	2.4	1.9	0.0	0.2
South Africa	74	93	116	25.3	27.2	1.4	4.6	0.4	41.1

**Table A2:** Traded Indices

This Table shows the list of traded funds used to construct net alpha, extending the approach of Berk and van Binsbergen (2015).

Ticker	CRSP Fund No	Fund Name	Asset Class	NAM	EUR	APA	EME	GLO
VFINX	31432	S&P 500 Index Fund	U.S. Large-Cap Blend	✓				✓
NAESX	31460	Small-Cap Index Fund	U.S. Small-Cap Blend	✓				✓
VEXMX	31433	Extended Market Index Fund	U.S. Mid and Small-Cap Blend	✓				
VIMSX	31473	Mid-Cap Index Fund	U.S. Mid-Cap Blend	✓				
VISGX	31471	Small-Cap Growth Index Fund	U.S. Small-Cap Growth	✓				
VISVX	31468	Small-Cap Value Index Fund	U.S. Small-Cap Value	✓				
VVIAX <sup>1</sup>	31457	Value Index Fund	U.S. Large-Cap Value	✓				✓
VFSVX	44222	All-World ex-US Small-Cap Index Fund	INT Small-Cap Blend		✓	✓	✓	✓
VGTSX	31200	Total International Stock Index Fund	INT Large-Cap Blend		✓	✓	✓	✓
VTRIX	31257	International Value Fund	INT Large-Cap Value		✓	✓	✓	✓
VEURX	31337	European Stock Index Fund	EUR Large-Cap Blend		✓			✓
VPACX	31336	Pacific Stock Index Fund	APA Large-Cap Blend			✓		✓
VEIEX	31338	Emerging Markets Stock Index Fund	EME Large-Cap Blend				✓	

<sup>1</sup> We use VIVAX (CRSP Fund No 031435) before VVIAX share class was introduced.



**Table A3:** The performance of funds sorted on eigenvalue centrality

This Table provides average monthly fund performance for portfolios conditional on eigenvalue centrality. At the end of each quarter, we sort funds into ten portfolios, based on eigenvalue centrality. Next, we track the equal-weighted performance of the portfolios during the next three months, when we rebalance. As fund performance measures, we use Net Alpha, Dollar Value Added, Gross Alpha, and DGTW returns. All performance measures are reported on a monthly basis and in percent, with the exception of Dollar Value Added which is reported in billion USD. We report time-series averages with t-statistics in parentheses based on Newey-West standard errors. Statistical significance at the 10%, 5%, and 1% level is indicated by \*, \*\*, and \*\*\*, respectively.

	Eigenvalue centrality										
	1 (low)	2	3	4	5	6	7	8	9	10 (high)	10 – 1
Net Alpha	0.079 (1.50)	0.083* (1.97)	0.019 (0.49)	-0.068** (-1.98)	-0.036 (-1.19)	-0.101*** (-3.50)	-0.092*** (-3.30)	-0.101*** (-4.04)	-0.120*** (-4.77)	-0.122*** (-5.25)	-0.201*** (-3.44)
Dollar Value Added	0.133 (0.30)	0.645** (2.12)	0.110 (0.35)	-0.142 (-0.42)	0.339 (0.90)	-0.444 (-0.91)	-0.413 (-0.92)	0.016 (0.03)	-0.621 (-1.07)	-1.876** (-2.18)	-2.009** (-1.99)
Gross Alpha	-0.080* (-1.89)	-0.052 (-1.23)	-0.057 (-1.34)	-0.129*** (-4.37)	-0.091*** (-2.61)	-0.110*** (-3.38)	-0.144*** (-5.07)	-0.156*** (-5.25)	-0.158*** (-5.54)	-0.159*** (-5.76)	-0.079*** (-2.94)
Gross DGTW	0.097 (1.58)	0.088 (1.09)	0.089 (1.33)	0.017 (0.3)	0.023 (0.5)	-0.008 (-0.14)	-0.025 (-0.68)	-0.022 (-0.74)	-0.049* (-1.66)	-0.051** (-2.09)	-0.148*** (-2.85)

**Table A4:** Fund centrality and characteristics

This table reports regression results for empirical models explaining variation in fund crowding

$$crowd_{i,t} = \beta \times \text{Covariates}_{i,t} + \alpha_i + \nu_t + \epsilon_{i,t}$$

where  $i$  denotes the mutual fund and  $t$  indexes time in quarters. The dependent variable for Models 1-4 is the baseline crowding measure and for Models 5 - 8 it is based on eigenvector centrality. The following variables are included in the vector  $\text{Covariates}_{i,t}$ :  $\log(\text{TNA})$  is the log of total net assets under management (in MM USD);  $\# \text{ Stocks}$  is the number of stocks in a fund portfolio;  $\# \text{ Countries}$  is the number of countries in which stock issuers are domiciled;  $\# \text{ Industries}$  is the number of ICB industries a fund invests in; Inverse normalized HFI (industries) is the inverse of the normalized Herfindahl Index on the ICB industry classification;  $\% \text{ North America}$ , for example, is the portfolio weight for investments in that region. Sample indicates whether either funds in the entire universe (full), domiciled in the United States (USA) or elsewhere (ex USA) are considered in the regression. Standard errors are double-clustered by fund and time, and the numbers in parantheses are p-values. Statistical significance at the 10%, 5%, and 1% level is indicated by \*, \*\*, and \*\*\*, respectively.

	Crowding				Eigenvector centrality			
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
$\log(\text{TNA})$	24.173*** (0.00)	27.645*** (0.00)	26.416*** (0.00)	28.169*** (0.00)	3.212*** (0.00)	3.203*** (0.00)	3.225*** (0.00)	2.875*** (0.00)
$\log(\text{TNA})^2$	2.148*** (0.00)	2.310*** (0.00)	2.840*** (0.00)	1.974*** (0.00)	0.290*** (0.00)	0.264*** (0.00)	0.360*** (0.00)	0.179*** (0.00)
$\# \text{ Stocks}$	0.054*** (0.00)	0.053*** (0.00)	0.050*** (0.00)	0.045*** (0.00)	0.006*** (0.00)	0.006*** (0.00)	0.005*** (0.00)	0.006*** (0.00)
$\# \text{ Countries}$	1.762*** (0.00)	1.782*** (0.00)	0.971*** (0.01)	1.616*** (0.00)	0.326*** (0.00)	0.262*** (0.00)	0.149*** (0.00)	0.238*** (0.00)
$\# \text{ Industries}$	8.623*** (0.00)	8.564*** (0.00)	7.719*** (0.00)	9.647*** (0.00)	0.879*** (0.00)	0.832*** (0.00)	0.865*** (0.00)	0.915*** (0.00)
Inverse normalized HFI (industries)	-0.000** (0.01)	-0.000** (0.03)	0.033 (0.74)	-0.000* (0.08)	-0.000*** (0.00)	-0.000** (0.04)	0.007 (0.61)	-0.000 (0.13)
$\log(\text{TNA}) \times \# \text{ Stocks}$	-0.018*** (0.00)	-0.017*** (0.00)	-0.010*** (0.00)	-0.030*** (0.00)	-0.002*** (0.00)	-0.002*** (0.00)	-0.001*** (0.00)	-0.004*** (0.00)
$\% \text{ North America}$	2.200*** (0.00)	2.143*** (0.00)	1.041*** (0.00)	2.580*** (0.00)	0.365*** (0.00)	0.405*** (0.00)	0.211*** (0.00)	0.482*** (0.00)
$\% \text{ Europe}$	2.219*** (0.00)	1.886*** (0.00)	3.149*** (0.00)	1.510*** (0.00)	0.292*** (0.00)	0.323*** (0.00)	0.428*** (0.00)	0.288*** (0.00)
$\% \text{ Japan}$	0.346 (0.33)	0.076 (0.74)	-0.935* (0.06)	0.195 (0.42)	-0.004 (0.89)	0.044 (0.14)	-0.192** (0.01)	0.095*** (0.00)
$\% \text{ Asia Pacific excluding Japan}$	-0.970*** (0.00)	-1.107*** (0.00)	-0.699* (0.08)	-1.274*** (0.00)	-0.027* (0.09)	0.004 (0.80)	-0.034 (0.40)	0.006 (0.73)
Sample	Full	Full	USA	ex USA	Full	Full	USA	ex USA
Number of funds	17,364	17,364	5,026	12,338	17,364	17,364	5,026	12,338
Number of observations	309,110	309,110	120,627	188,483	309,110	309,110	120,627	188,483
$R^2$	0.184	0.167	0.090	0.275	0.212	0.206	0.095	0.335
Fund Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes

**Table A5:** The performance of funds sorted on crowding, using value-weighting of funds or factor models

This Table provides average monthly fund performance for portfolios conditional on crowding. At the end of each quarter, we sort funds into ten portfolios, based on crowding. In Panel A, we track the value-weighted performance of the portfolios during the next three months, when we rebalance. As fund performance measures, we use Net Alpha and Gross Alpha. All performance measures are reported on a monthly basis and in percent, with the exception of Dollar Value Added which is reported in billion USD. We report time-series averages with t-statistics in parentheses based on Newey-West standard errors. In Panel B, we track the equal-weighted excess return of the portfolios during the next three months, when we rebalance. We obtain a time-series of portfolio returns. We use different factor models to adjust for risk. We report estimated alpha with t-statistics in parentheses. All performance measures are reported on a monthly basis and in percent, with the exception of Dollar Value Added which is reported in billion USD. We report time-series averages with t-statistics in parentheses based on Newey-West standard errors. Statistical significance at the 10%, 5%, and 1% level is indicated by \*, \*\*, and \*\*\*, respectively.

<b>Panel A:</b> Value-weighted portfolio returns											
	1 (low)	2	3	4	5	Crowding		8	9	10 (high)	10 – 1
						6	7				
Net Alpha	0.049 (0.82)	0.070** (2.21)	0.001 (0.01)	0.024 (0.68)	0.034 (1.08)	−0.043 (−1.25)	−0.007 (−0.30)	−0.016 (−0.64)	−0.054** (−1.99)	−0.087** (−2.41)	−0.136* (−1.83)
Gross Alpha	−0.032 (−0.59)	−0.040 (−1.13)	−0.075* (−1.73)	−0.039 (−0.93)	−0.032 (−0.99)	−0.116*** (−3.02)	−0.078** (−2.08)	−0.119*** (−3.60)	−0.132*** (−3.96)	−0.154*** (−4.41)	−0.123*** (−2.61)
<b>Panel B:</b> Factor model alphas											
	1	2	3	4	5	Crowding		8	9	10	10 – 1
						6	7				
CAPM	0.139 (1.16)	0.038 (0.37)	−0.000 (−0.00)	−0.157 (−1.58)	−0.171** (−2.57)	−0.203*** (−3.56)	−0.173*** (−3.79)	−0.200*** (−4.04)	−0.240*** (−3.88)	−0.233*** (−3.80)	−0.372** (−2.46)
FF 3-Factor	−0.004 (−0.05)	−0.020 (−0.22)	−0.017 (−0.16)	−0.150 (−1.56)	−0.115* (−1.69)	−0.125*** (−2.94)	−0.128*** (−3.27)	−0.155*** (−4.10)	−0.182*** (−3.96)	−0.160*** (−4.01)	−0.156** (−1.99)
Carhart 4-Factor	0.011 (0.14)	−0.013 (−0.14)	−0.040 (−0.34)	−0.165 (−1.63)	−0.103 (−1.60)	−0.113*** (−2.75)	−0.130*** (−3.14)	−0.163*** (−4.21)	−0.193*** (−4.09)	−0.161*** (−3.92)	−0.173** (−2.13)
FF 5-Factor	0.002 (0.03)	0.057 (0.66)	0.006 (0.05)	−0.169* (−1.68)	−0.092 (−1.38)	−0.066 (−1.49)	−0.107** (−2.46)	−0.152*** (−3.63)	−0.184*** (−3.53)	−0.143*** (−3.08)	−0.145* (−1.69)

**Table A6:** The performance of funds in sequential sorts on size and crowding, using alternative performance measures and eigenvalue centrality

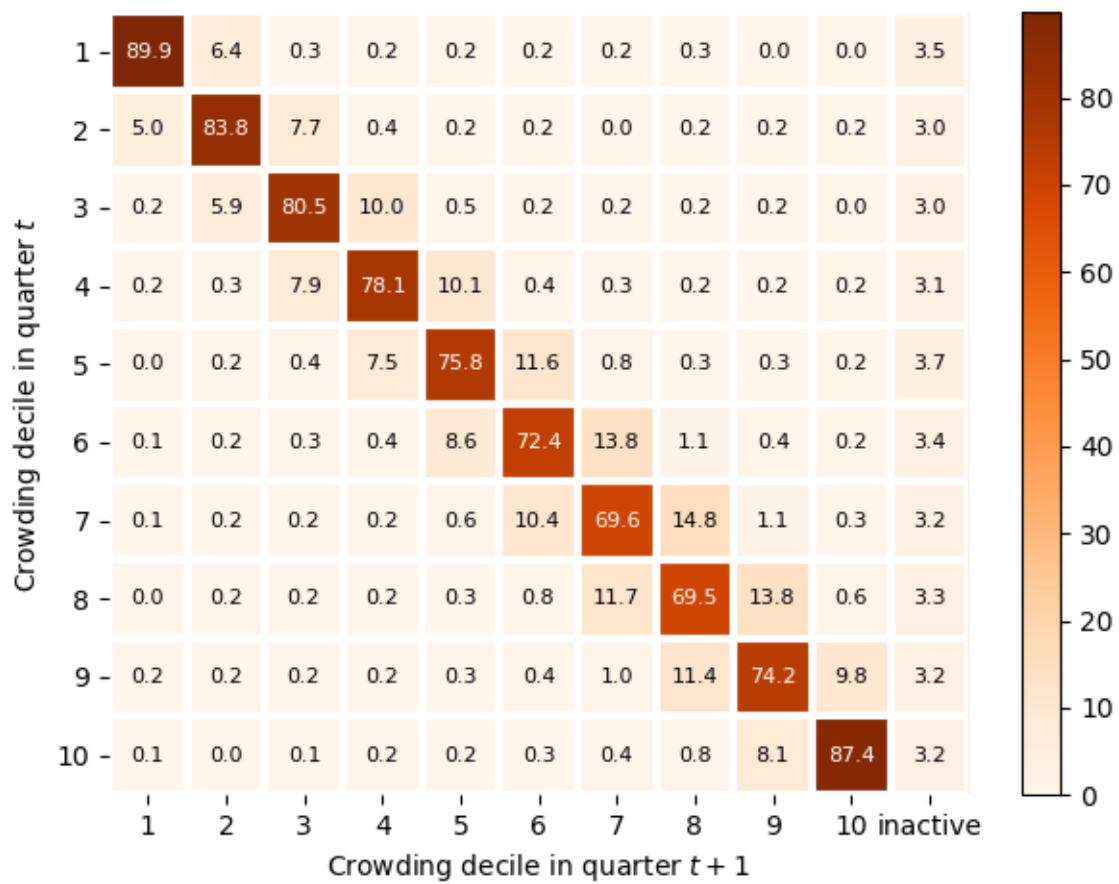
This Table provides average monthly fund performance for portfolios conditional on size and crowding. At the end of each quarter, we sort funds into ten portfolios, based on size. Next, we sort funds within each size decile into three portfolios based on our baseline measure of crowding (Panel A) or eigenvalue centrality (Panel B). Next, we track the equal-weighted performance of the portfolios during the next three months, when we rebalance. As fund performance measures, we use Dollar Value Added, Gross Alpha, and DGTW returns in Panel A and B, and additionally Net Alpha in Panel B. All performance measures are reported on a monthly basis and in percent, with the exception of Dollar Value Added which is reported in billion USD. We report time-series averages with t-statistics in parentheses based on Newey-West standard errors. Statistical significance at the 10%, 5%, and 1% level is indicated by \*, \*\*, and \*\*\*, respectively.

<b>Panel A: Crowding</b>										
	1 (low)	2	3	4	5	6	7	8	9	10 (high)
Dollar Value Added	-0.030* (-1.67)	-0.044** (-2.15)	-0.112*** (-3.38)	-0.134** (-2.25)	-0.204** (-2.22)	-0.210 (-1.52)	-0.454** (-2.49)	-1.009*** (-2.80)	-1.108** (-2.27)	-7.201* (-1.93)
Gross Alpha	-0.020 (-0.52)	-0.093*** (-3.90)	-0.098*** (-4.00)	-0.099*** (-3.78)	-0.064** (-2.27)	-0.080*** (-3.62)	-0.108*** (-3.94)	-0.109*** (-3.82)	-0.104*** (-3.97)	-0.171*** (-5.09)
Gross DGTW	-0.025 (-0.51)	-0.161*** (-3.36)	-0.125** (-2.41)	-0.161*** (-3.31)	-0.133*** (-2.66)	-0.129*** (-2.84)	-0.174*** (-3.59)	-0.168*** (-3.56)	-0.148*** (-2.97)	-0.214*** (-3.99)
<b>Panel B: Eigenvector centrality</b>										
	1 (low)	2	3	4	5	6	7	8	9	10 (high)
Net Alpha	-0.162*** (-3.13)	-0.156*** (-3.71)	-0.194*** (-4.37)	-0.162*** (-3.23)	-0.150*** (-2.88)	-0.150*** (-2.83)	-0.188*** (-4.05)	-0.193*** (-3.88)	-0.141*** (-3.92)	-0.176*** (-5.64)
Dollar Value Added	-0.036** (-2.40)	-0.052** (-2.56)	-0.114*** (-3.33)	-0.137** (-2.24)	-0.200** (-2.08)	-0.259* (-1.82)	-0.568*** (-2.98)	-1.074*** (-3.00)	-1.224** (-3.00)	-7.448** (-2.04)
Gross Alpha	-0.064 (-1.51)	-0.089*** (-3.36)	-0.099*** (-3.67)	-0.096*** (-3.34)	-0.064** (-2.05)	-0.086*** (-3.11)	-0.111*** (-3.41)	-0.112*** (-3.41)	-0.104*** (-3.59)	-0.155*** (-4.44)
Gross DGTW	-0.061 (-1.30)	-0.149*** (-3.06)	-0.127** (-2.49)	-0.155*** (-3.09)	-0.137** (-2.57)	-0.127*** (-2.62)	-0.168*** (-3.44)	-0.164*** (-3.17)	-0.139*** (-2.71)	-0.197*** (-3.61)

**Table A7:** The performance of funds sorted on crowding, using an asymmetric network

This Table provides average monthly fund performance for portfolios conditional on crowding, defined using an asymmetric network instead of portfolio holdings overlap. At the end of each quarter, we sort funds into ten portfolios, based on crowding (Panel A) or eigenvalue centrality (Panel B). Next, we track the equal-weighted DGTW returns of the portfolios during the next three months, when we rebalance. The performance measure is reported on a monthly basis and in percent. We report time-series averages with t-statistics in parentheses based on Newey-West standard errors. Statistical significance at the 10%, 5%, and 1% level is indicated by \*, \*\*, and \*\*\*, respectively.

<b>Panel A: Crowding</b>										
1 (low)	2	3	4	5	6	7	8	9	10 (high)	10 – 1
0.096	0.081	0.103	0.049	0.036	-0.026	-0.020	-0.020	-0.044	-0.078***	-0.173***
(0.164)	(0.182)	(0.122)	(0.429)	(0.447)	(0.569)	(0.575)	(0.505)	(0.107)	(0.005)	(0.005)
<b>Panel B: Eigenvector centrality</b>										
1 (low)	2	3	4	5	6	7	8	9	10 (high)	10 – 1
0.085	0.124	0.069	0.043	0.013	-0.002	-0.020	-0.021	-0.028	-0.085***	-0.170***
(0.177)	(0.108)	(0.328)	(0.428)	(0.790)	(0.964)	(0.609)	(0.499)	(0.287)	(0.004)	(0.002)



**Figure 1:** Persistence of crowding

Note: Conditional on any decile in quarter  $t$ , this figure plots the probability of transitioning to decile  $j$  or going inactive at quarter  $t + 1$ . The time period of analysis is between 2001 Q3 and 2014 Q1 (51 quarters).