What Alleviates Crowding in Factor Investing?*

Victor DeMiguel Alberto Martín-Utrera Raman Uppal

December 8, 2020

Abstract

Factor investing is a low-cost approach to active fund management that exploits common firm characteristics. The growing number of institutions exploiting these strategies raises concerns that *crowding* may increase price-impact costs and erode profits. We identify a mechanism that alleviates crowding in factor investing—*trading diversification*: institutions exploiting different characteristics can reduce each other's price-impact costs *even when* their rebalancing trades are not negatively correlated. We study how trading diversification affects the equilibrium and find that, while competition to exploit a characteristic erodes its profits because of crowding, competition among institutions exploiting *other* characteristics alleviates crowding. Empirically, we find that trading diversification can increase the optimal investment positions by around 50% and the optimal profits by around 25%.

Keywords: capacity of quantitative strategies, price impact, competition.
JEL Classification: G11, G12, G23, L11.

*DeMiguel, London Business School, e-mail: avmiguel@london.edu; Martín-Utrera, New Jersey Institute of Technology, e-mail: amarut@njit.edu; Uppal, EDHEC Business School, EDHEC Scientific Beta Research Chair, and CEPR, e-mail: raman.uppal@edhec.edu. We are grateful for comments from Noel Amenc, Simcha Barkai, Jean-Pierre Benoît, Jonathan Berk, Svetlana Bryzgalova, Laurent Calvet, João Cocco, Thomas Dangl, James Dow, David Feldman (discussant), Gavin Feng, Francesco Franzoni, Thierry Foucault, Felix Goltz, Francisco Gomes, Amit Goyal, Christopher Hennessy, Nicholas Hirschey, Hugues Langlois (discussant), Christian Heyerdahl-Larsen, Bruce Jacobs, Apostolos Kourtis, Augustin Landier, Ashley Lester, Ken Levy, Abraham Lioui, Christian Lundblad (discussant), Raphael Markellos, Narayan Naik, Vasant Naik, Emre Ozdenoren, Anna Pavlova, Scott Richardson, Vania Stavrakeva, Nikos Tessaromatis, David Thesmar, Cristian Tiu (discussant), Afonso Varatojo, Josef Zechner, and seminar participants at the 12th Annual Hedge Fund Research Conference, 27th Finance Forum, 2020 INFORMS Annual Meeting, 2020 NFA Conference, 4th SAFE Market Microstructure Conference, EDHEC Business School, INSEAD, King's Business School, London Business School, Norwich Business School, SKEMA Business School, Schroeders, SYZ Asset Management, Université Catholique de Louvain, and University of St. Gallen.

1 Introduction

Factor investing is a low-cost approach to active fund management that exploits common characteristics, such as value, investment, and profitability.[1] Although assets under management in factor-investing strategies have grown rapidly at 30% per annum since 2010, reaching $700 billion by 2018 (Ratcliffe, Miranda, and Ang, 2017; Riding, 2018), their capacity remains limited by price-impact costs. Indeed, there is a large literature that characterizes a strategy's capacity, defined as the total investment that can be allocated to it before price-impact costs erode its profits entirely.[2] Importantly, the growth in factor investing has been accompanied by an explosion in the *number* of institutions exploiting these strategies. For instance, 145 managers launched factor-investing products in 2018 (Flood, 2019). This raises concerns about *crowding*: as an increasing number of institutions exploit the same characteristic, competition leads them to overinvest and price-impact costs erode profits.[3]

Our first contribution is to identify a mechanism that alleviates crowding in factor investing—*trading diversification*: institutions exploiting *different* characteristics can reduce each other's price-impact costs. It is intuitive that institutions exploiting different characteristics whose portfolio-rebalancing trades are negatively correlated reduce each other's price-impact costs because their trades net out on average. However, we show theoretically that combining characteristics may reduce price-impact costs *even when* their rebalancing trades are not negatively correlated. Empirically, we consider 18 characteristics and find that there is a reduction of around 17% in price-impact costs when considering them in combination, relative to the cost of trading them in isolation. More importantly, when the characteristics are exploited in combination, the total capacity and total optimal investment

[1] In the investment industry, factor investing strategies are often referred to as smart beta.

[2] See, for instance, Korajczyk and Sadka (2004); Novy-Marx and Velikov (2016); Ratcliffe et al. (2017) and Frazzini, Israel, and Moskowitz (2018).

[3] For instance, in his AFA presidential address Stein (2009) argues that "basic economic logic suggests that as more money is brought to bear against a given trading opportunity, any predictable excess returns must be reduced and eventually eliminated." Similarly, Jacobs and Levy (2014) state that: "Smart beta strategies are often based on common, generic factors used by many managers. This approach leaves their performance susceptible to factor crowding: Too many investors are buying (or selling) the same securities on the basis of the same factors."

across the 18 characteristics increase by around 50% and the total profits increase by around 25%, relative to exploiting them in isolation.

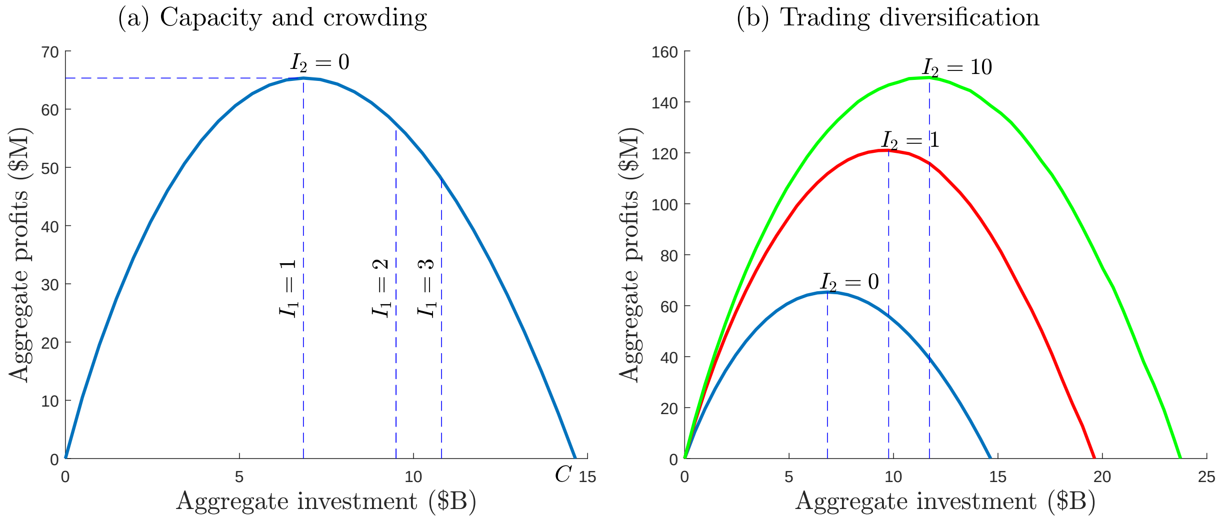
Our second contribution is to study how trading diversification affects equilibrium investment positions and profits in the factor-investing industry. To do this, we develop a game-theoretic model related to those by Berk and Green (2004) and Pástor and Stambaugh (2012), who consider competition among fund managers facing diseconomies of scale at the *fund* and *industry* levels, respectively. In contrast, we consider a model with two groups of factor investors, each group exploiting a different characteristic. Investors within each group compete to exploit the same characteristic and thus face diseconomies of scale at the *characteristic* level driven by price-impact costs. However, there is a *positive externality* between the two groups of investors because they reduce each other's price-impact costs due to trading diversification across characteristics.

We characterize in closed form the equilibrium investment positions and profits in the game and study how trading diversification affects them. To gauge the magnitude of the effect, we also calibrate the model using "investment (asset growth)" as the first characteristic and "gross profitability" as the second. Our main findings are illustrated in Figure 1. Panel (a) illustrates the effect of crowding by depicting the aggregate profits of the investors exploiting the first characteristic as a function of their aggregate investment, when there are no investors exploiting the second characteristic. The graph shows that when there is only a single investor exploiting the first characteristic ($I_1 = 1$), she maximizes her profits by investing around *half* of the characteristic's capacity C. However, as the number of investors in the first characteristic I_1 increases, competition leads them to overinvest and, as a result, their aggregate profits decrease. In the limit, as the number of investors goes to infinity, we find that their aggregate investment position converges to the strategy's capacity and their aggregate profits converge to zero because of price-impact costs. Thus, *competition among investors exploiting the same characteristic erodes their profits because of crowding.*

Panel (b) of Figure 1 illustrates the effect of trading diversification and competition among investors exploiting the second characteristic. The graph compares the aggregate

Figure 1: Crowding and trading diversification

This figure illustrates the effect of crowding and trading diversification on investment positions and profits. Panel (a) illustrates the effect of crowding by depicting the aggregate profits of the investors exploiting the first characteristic as a function of their aggregate investment, when there are no investors exploiting the second characteristic ($I_2 = 0$). The graph also shows the optimal aggregate investment in the first characteristic when there are $I_1 = 1, 2, 3$ investors exploiting it. Panel (b) illustrates the effect of trading diversification and competition among investors exploiting the second characteristic by comparing the aggregate profits of investors exploiting the first characteristic for the cases where: (i) there are no investors exploiting the second characteristic ($I_2 = 0$) and thus there is no trading diversification, (ii) there is one investor ($I_2 = 1$) and (iii) there are ten investors ($I_2 = 10$) exploiting the second characteristic. The figure is calibrated using "investment (asset growth)" as the first characteristic and "gross profitability" as the second.



profits of investors exploiting the first characteristic for the cases where: (i) there are no investors exploiting the second characteristic ($I_2 = 0$) and thus there is no trading diversification, (ii) there is one investor ($I_2 = 1$) and (iii) there are ten investors ($I_2 = 10$) exploiting the second characteristic. Comparing the case of $I_2 = 0$ with that of $I_2 = 1$, we observe that trading diversification increases the capacity of the first characteristic as well as its equilibrium aggregate investment position and profits. Thus, *trading diversification alleviates crowding in factor investing.* Moreover, an increase in the number of investors exploiting the second characteristic further increases the capacity, aggregate investment position, and aggregate profits associated with the first characteristic. *Thus, competition among investors exploiting other characteristics further alleviates crowding.*

Our work has implications for the industrial organization and regulation of the quantitative investment industry. First, financial institutions should focus not only on characteristics that are profitable, but also are exploited by a *small number* of institutions. This prediction of our model is reflected in the current structure of the investment industry, with three institutions—BlackRock, Vanguard, and State Street—holding 79% of the assets in ETF products (Baert, 2018). Moreover, institutions can increase their market power by increasing assets under management or acquiring competitors, a strategy recently adopted by Invesco to become the fourth largest ETF provider in the U.S. (Carlson, 2019). Second, financial institutions should exploit characteristics that allow them to benefit from trading diversification. For instance, we show that the institutions exploiting an "investment (asset growth)" characteristic benefit from the trading diversification generated by other institutions exploiting "gross profitability". Similarly, Frazzini, Israel, and Moskowitz (2015) find using proprietary data that "value and momentum trades tend to offset each other, resulting in lower turnover which has real transaction costs benefits." Third, regulators need to recognize that, although encouraging competition among fund managers exploiting a characteristic may reduce fees (Wahal and Wang, 2011), it may also erode fund returns because of crowding. However, encouraging the *appropriate* balance of competition between managers exploiting *different* characteristics can actually alleviate crowding and increase profits due to trading diversification.

Our work is closely related to Bonelli, Landier, Simon, and Thesmar (2019), who consider competitive traders who exploit a single investment signal and maximize multiperiod mean-variance utilities. They analyze how the capacity and performance of the strategy depend on the *persistence* of the signal and the traders' estimate of the number of competitors. In contrast, we focus on how *trading diversification* affects capacity and performance when there is competition among investors exploiting *different* characteristics.

There is also a literature on competition in the active mutual-fund industry; see the review by Berk and van Binsbergen (2017). The seminal paper by Berk and Green (2004) considers managers who have different abilities to generate alpha and face diseconomies of scale at the *fund* level. In contrast, Pástor and Stambaugh (2012) assume diseconomies of scale at the *industry* level and Pástor, Stambaugh, and Taylor (2015) provide empirical evidence supporting this assumption.[4] We consider diseconomies of scale at the *characteristic* level, but we provide a microfoundation for them based on price-impact costs, which we estimate from data on firm characteristics. This microfoundation is consistent with the empirical evidence in Edelen, Evans, and Kadlec (2007) and Pástor et al. (2015) that suggests trading costs are the primary source of diseconomies of scale. A key feature that distinguishes our work from these papers is that we consider competition among investors exploiting *different* characteristics, which *alleviates* the diseconomies of scale because of trading diversification. Harvey, Liu, Tan, and Zhu (2020) also study crowding in investment management, but they focus on the impact of team management on discretionary funds, whereas we focus on the impact of trading diversification on factor-investing funds.

Our work is also related to the literature on the capacity of quantitative strategies. Several papers study the capacity of strategies that exploit a *single* characteristic: Korajczyk and Sadka (2004) study the market-impact costs associated with exploiting momentum and find that this characteristic can be exploited on only a relatively modest scale. Novy-Marx and Velikov (2016) consider 23 anomalies and find that simple strategies to mitigate transaction costs significantly reduce price impact and thus increase the scale to which the characteristics can be exploited. The aforementioned papers use publicly available datasets to estimate the trading costs of an average investor. In contrast, Ratcliffe et al. (2017) and Frazzini et al. (2018) use proprietary data from large money managers and find that the trading costs of these financial institutions are quite small, and thus, they can exploit these characteristics to a much larger extent than previously thought. We build on these papers by showing that trading diversification across characteristics can further increase capacity as well as the equilibrium investment positions and profits of factor investors.

Other papers have also found that combining characteristics helps to reduce transaction costs. For instance, Barroso and Santa-Clara (2015) consider currency portfolios

[4] In addition, Feldman, Saxena, and Xu (2018, 2019) show that when industry concentration is lower, net alpha and industry size are smaller. Other notable work in this area include Wahal and Wang (2011), who find that incumbent funds that have high overlap in holdings with entrant funds reduce management fees and suffer lower alphas, and Hoberg, Kumar, and Prabhala (2018, 2019), who show that buy-side competition among mutual funds explains future alphas.

based on six characteristics and explain that "transaction costs depend crucially on the time-varying interaction between characteristics." Novy-Marx and Velikov (2016) study "filtering," a cost mitigation technique that allows investors trading one strategy to opportunistically take small positions in another at effectively negative trading costs. Frazzini, Israel, and Moskowitz (2015) show that value and momentum trades offset each other. DeMiguel, Martin-Utrera, Nogales, and Uppal (2020) show that transaction costs increase the dimension of the cross-section of stock returns because "combining characteristics allows one to diversify trading, just as combining them allows one to diversify risk." Our work differs from these papers because we show how the *strategic interactions* among financial institutions can alleviate crowding concerns in factor investing due to trading diversification.

The remainder of this manuscript is organized as follows. Section 2 describes how we extend the parametric portfolios of Brandt, Santa-Clara, and Valkanov (2009) to consider price-impact costs. Section 3 analyzes trading diversification theoretically and empirically. Section 4 develops the game-theoretic model and characterizes its equilibrium. Section 5 discusses the effect of the strategic interactions among investors on investment positions and profits. Section 6 takes our model to the data. Section 7 concludes. Appendix A provides the proofs for all results and the Internet Appendix contains robustness checks.

2 Parametric portfolios with price-impact costs

In this section, we explain how we extend the parametric portfolios of Brandt et al. (2009) to consider price-impact costs in factor investing. We rely on the resulting framework for our theoretical and empirical analysis in the remainder of the manuscript.

2.1 Parametric portfolios

We consider a market with N stocks and K firm-specific characteristics. Let $r_t \in \mathbb{R}^N$ be the stock-return vector and $x_{kt} \in \mathbb{R}^N$ the kth-characteristic portfolio-weight vector at time t. For instance, x_{1t} could contain a portfolio based on the "investment (asset growth)" of the

N firms and x_{2t} based on their "gross profitability." We standardize x_{kt} cross sectionally so that it has zero mean; that is, x_{kt} is a long-short portfolio, and thus, has zero cost. This is customary in cross-sectional asset pricing and it facilitates our analysis by removing the need for a budget constraint.[5] We also standardize the characteristic portfolio weight vectors so that the sum of the positive or negative weights is one; that is, a portfolio $\theta_k x_{kt}$ invests θ_k dollars on both the positive and negative legs. Finally, for the empirical analysis we consider *value-weighted* long-short characteristic portfolios that do not allocate large weights to small firms that are difficult to trade.

Like Brandt et al. (2009), we consider a parametric portfolio policy such that the weight on a particular stock at time t is a linear function of *only* its weights in the K characteristic portfolios. Moreover, the same linear function is applied across stocks and through time.[6] Thus, the parametric portfolio at time t can be written as

$$w_t(\theta) = \sum_k x_{kt}\theta_k = X_t\theta, \tag{1}$$

where $\theta_k \in \mathbb{R}$ is the investment position in the kth characteristic, $\theta = (\theta_1, \theta_2, \ldots, \theta_K)$ is the investment-position vector, and $X_t = (x_{1t}, x_{2t}, \ldots, x_{Kt}) \in \mathbb{R}^{N\times K}$ is the matrix whose columns are the K long-short characteristic portfolios at time t. The return of the parametric portfolio at time $t+1$ is:[7]

$$r_{p,t+1} = w_t(\theta)^\top r_{t+1} = \sum_k \theta_k\, x_{kt}^\top r_{t+1} = \theta^\top X_t^\top r_{t+1}. \tag{2}$$

[5]There are both long-short and long-only factor-investing products in financial markets. The advantages of long-short funds are that they are market neutral and they can exploit the favorable performance of the short leg. The main advantage of long-only products is that they do not require shorting, which may be costly. However, many long-short products can reduce costs by shorting the market index instead of individual stocks.

[6]The assumption that the weights assigned to the characteristics are *constant* over time is consistent with the empirical literature on the capacity of quantitative strategies, which characterizes the largest investment position that can be allocated to a particular characteristic over the *entire* period before price-impact costs drive its net average return to zero; see Korajczyk and Sadka (2004), Ratcliffe et al. (2017), and Frazzini et al. (2018).

[7]Although r_{t+1} is a *payoff* because the parametric portfolio is a zero-cost long-short portfolio, for simplicity we refer to it as a *return*.

2.2 Price-impact costs

Financial institutions trade for both informational and liquidity motives; see, for instance, Grossman and Stiglitz (1980) and Kyle (1985). While informational trades result in permanent price impact, liquidity trades have temporary price impact. Given our focus on factor-investing strategies that exploit *publicly known* characteristics, we focus on temporary price-impact costs, which is also consistent with the empirical literature on quantitative-strategy capacity. However, in unreported results we observe that our findings are robust to considering persistent price-impact costs.

While several papers assume the price impact of a trade is linear in the amount traded (Korajczyk and Sadka, 2004; Novy-Marx and Velikov, 2016), empirical evidence finds that price impact grows with the square root of the amount traded (Torre and Ferrari, 1997; Grinold and Kahn, 2000; Almgren, Thum, Hauptmann, and Li, 2005; Ratcliffe et al., 2017; Frazzini et al., 2018). To capture both specifications, we write the general *price-impact function* at time t as:

$$\mathrm{PI}_t = \Lambda_t \ \mathrm{sign}(\Delta w_t) \circ |\Delta w_t|^{\alpha}, \tag{3}$$

where the case with $\alpha = 1$ corresponds to a linear price-impact function and the case with $\alpha = 1/2$ to the square-root price-impact function, and where

$$\Lambda_t = \mathrm{diag}(\lambda_{t1}, \lambda_{t2}, \ldots, \lambda_{tN}) \in \mathbb{R}^{N\times N} \tag{4}$$

is the diagonal matrix whose nth element, λ_{tn}, is the price-impact parameter for the nth stock at time t, which is exogenous in our model; $\Delta w_t \in \mathbb{R}^N$ is the *aggregate-trade vector* at time t, defined as the vector that contains the net total amount traded in the market for each stock and given by

$$\Delta w_t = \sum_k \theta_k \tilde{x}_{kt} \quad \text{where} \tag{5}$$

$$\tilde{x}_{kt} = x_{kt} - x_{k,t-1} \circ (e + r_t), \tag{6}$$

$x \circ y$ is the componentwise or Hadamard product of vectors x and y, e is the N-dimensional vector of ones, and $\mathrm{sign}(x)$, $|x|$, and x^{α} are the componentwise sign, absolute value, and power function of vector x, respectively. The aggregate-trade vector can also be conveniently written in matrix notation as

$$\Delta w_t = \tilde{X}\theta, \tag{7}$$

where $\tilde{X}_t = (\tilde{x}_{1t}, \tilde{x}_{2t}, \ldots, \tilde{x}_{Kt}) \in \mathbb{R}^{N \times K}$.

The *price-impact cost* at time t is the amount of trading multiplied by its price impact:

$$\text{PIC}_t = \Delta w_t^\top \ \text{PI}_t = \Delta w_t^\top \ \Lambda_t \ \text{sign}(\Delta w_t) \circ |\Delta w_t|^\alpha. \tag{8}$$

Then, substituting (7) into (8), the price-impact cost of rebalancing the portfolio at time t can be written as

$$\text{PIC}_t = \theta^\top \tilde{X}_t^\top \ \Lambda_t \ (\text{sign}(\tilde{X}_t\theta) \circ |\tilde{X}_t\theta|^\alpha). \tag{9}$$

2.3 Optimal parametric portfolio

The optimal portfolio at time t is given by the investment-position vector θ that optimizes the *conditional* expected portfolio return net of price-impact costs. However, a key insight of Brandt et al. (2009) is that the optimal parametric portfolio policy can be obtained by optimizing the *unconditional* expectation because the investment-position vector θ is assumed to be *constant* through time. In addition, for simplicity we assume that investors are risk neutral, although Appendix IA.1.1 shows that our results are robust to considering risk-averse investors. Therefore, the optimal parametric portfolio is obtained by choosing the investment-position vector θ that optimizes the unconditional expectation of the difference between the price-impact cost and the return

$$\min_{\theta} \quad E\big[\text{PIC}_t - r_{p,t+1}\big], \tag{10}$$

in which the portfolio return $r_{p,t+1}$ and the price-impact cost PIC_t are functions of θ, as specified in Equations (2) and (9), respectively.

3 Trading diversification

In this section, we characterize the trading-diversification mechanism from both a theoretical and an empirical perspective. To do this, we study the effect on price-impact costs of *combining* characteristics.

3.1 Theoretical results

We first study theoretically how the price-impact costs of exploiting characteristics are reduced when they are traded in combination.

Definition 3.1 Given K characteristics whose rebalancing trades follow a particular joint probability distribution, the *price-impact diversification ratio* for the nth stock is defined as the ratio of the unconditional expected price-impact cost required to rebalance the position on the nth stock for an equally weighted portfolio of the K characteristics to that required to rebalance the two characteristics in isolation; that is

$$\text{price-impact diversification ratio} := \frac{E\left[\lambda_{tn}\left|\sum_{k=1}^{K} \tilde{x}_{ktn}\right|^{1+\alpha}\right]}{\sum_{k=1}^{K} E\left[\lambda_{tn}\left|\tilde{x}_{ktn}\right|^{1+\alpha}\right]}, \tag{11}$$

where λ_{tn} is the nth stock price-impact parameter at time t, that is, the nth element of the diagonal matrix Λ_t in (4), and $\tilde{x}_{ktn}$ is the trade on the nth stock required to rebalance the kth characteristic at time t, that is, the nth element of vector $\tilde{x}_{kt}$ in (6).

Note that a price-impact diversification ratio smaller than one implies that there is a reduction in price-impact costs from combining characteristics. For instance, a price-impact diversification ratio of 0.75 would indicate the price-impact cost of trading the characteristics in combination is 25% smaller than that of trading them in isolation. On the other hand, a price-impact diversification ratio of one implies that there are no diversification benefits from combining the characteristics. Finally, a price-impact diversification ratio larger than one implies that the price-impact cost of trading the characteristics in combination is higher than that of trading them in isolation.[8]

The following proposition characterizes analytically the price-impact diversification ratio for $\alpha > -1$. DeMiguel et al. (2020, Proposition 3) characterize this ratio for the case with proportional transaction costs, $\alpha = 0$. Here, we generalize their result to the case with $\alpha > -1$, which includes the two relevant cases: (i) $\alpha = 1$, implying a linear price-impact function, and thus, quadratic price-impact costs and (ii) $\alpha = 0.5$, implying a square-root price-impact function, and thus, subquadratic price-impact costs.

Proposition 3.1 *Assume that the trades in the nth stock required to rebalance K characteristics, that is, the quantities $\tilde{x}_{ktn}$ for $k = 1, 2, \ldots, K$, are jointly distributed as a Normal distribution with zero mean and positive definite covariance matrix Ω. Moreover, assume the nth stock price-impact parameter is independently distributed from the rebalancing trades. Then, for any $\alpha > -1$ the*

$$\textit{price-impact diversification ratio} = \frac{\left(\sum_{k=1}^{K} \sigma_k^2 + \sum_{k=1}^{K} \sum_{l \neq k} \rho_{kl} \sigma_k \sigma_l\right)^{\frac{1+\alpha}{2}}}{\sum_{k=1}^{K} \sigma_k^{1+\alpha}},$$

where σ_k^2 is the variance of the rebalancing trade $\tilde{x}_{ktn}$ and ρ_{kl} is the correlation between the rebalancing trades $\tilde{x}_{ktn}$ and $\tilde{x}_{ltn}$. If, in addition, the covariance matrix Ω is symmetric with respect to the K characteristics, that is, if $\sigma_k^2 = \sigma^2$ for all k and $\rho_{kl} = \rho$ for all $k \neq l$, then[9]

$$\textit{price-impact diversification ratio} = \frac{\left[K(1 + (K-1)\rho)\right]^{\frac{1+\alpha}{2}}}{K} \tag{12}$$

and the price-impact diversification ratio is strictly smaller than one if and only if

$$\rho < \bar{\rho} = \frac{K^{\frac{1-\alpha}{1+\alpha}} - 1}{K - 1}. \tag{13}$$

[8] Note that in practice it is not feasible to trade characteristics in isolation because they all require trading in the same underlying stocks. However, if the rebalancing trades of K characteristics are highly positively correlated, then if one could trade them in K isolated markets (that is, trades in each market do not affect prices in the other $K-1$ markets), then the price-impact cost would be smaller than that of trading them in the same market because price-impact costs are a strictly convex function of the amount traded.

[9] Note that the term $K(1 + (K-1)\rho)$ is strictly positive because Ω is positive definite.

We now discuss Proposition 3.1 focusing, for simplicity, on the symmetric case with $\sigma_k^2 = \sigma^2$ for all k and $\rho_{kl} = \rho$ for all $k \neq l$. Equation (13) shows that for the case with a linear price-impact function, $\alpha = 1$, we have that $\bar{\rho} = 0$, and thus the price-impact diversification ratio is smaller than one only when the correlation between the rebalancing trades of the different characteristics is negative, $\rho < 0$. This is what intuition would suggest: combining characteristics whose portfolio-rebalancing trades are negatively correlated reduces price-impact costs because their trades net out on average.

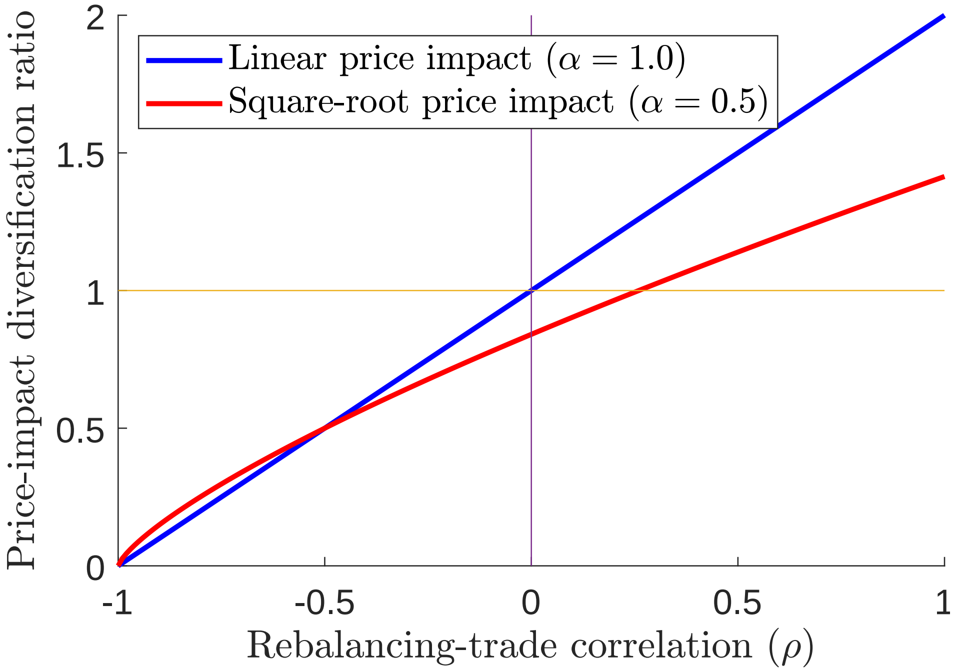
The surprising result is for the empirically relevant case of square-root price-impact function, $\alpha = 0.5$. In this case, the price-impact diversification ratio can be smaller than one *even* when the rebalancing trades are positively correlated, as long as the correlation is below $\bar{\rho} = (K^{1/3} - 1)/(K - 1) > 0$. That is, with square-root price impact, combining characteristics leads to a reduction of price-impact costs *even* when the rebalancing trades of the two characteristics are moderately *positively* correlated.

These theoretical findings are illustrated in Figure 2. Panel (a) depicts how the price-impact diversification ratio depends on the rebalancing trade correlation ρ for the case with two characteristics and linear or square-root price impact ($\alpha = 0.5$). The graph shows that the ratio is smaller than one provided $\rho < \bar{\rho} = 2^{0.5/1.5} - 1 \approx 0.26$. For example, if $\rho = 0$, the price-impact diversification ratio for the case with two characteristics is around 0.84, indicating that the price-impact cost of trading the two characteristics in combination is around 16% smaller than the average cost of trading them in isolation. Panel (b) illustrates the effect of the number of characteristics K on the price-impact diversification ratio for the case with zero rebalancing trade correlation ($\rho = 0$) and square-root price impact ($\alpha = 1/2$). The plot shows that the impact of trading diversification increases substantially with the number of characteristics combined and the price-impact diversification ratio ranges from 84% for the case with two characteristics to around 50% for the case with $K = 18$ characteristics.

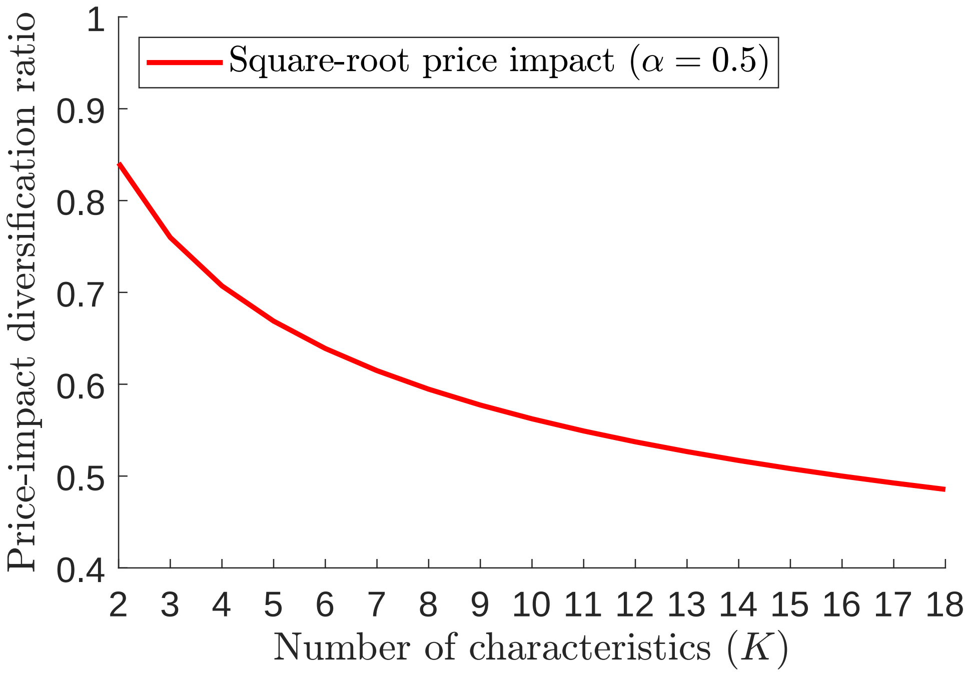
Figure 2: Price-impact diversification ratio: Theoretical results

This figure depicts the price-impact diversification ratio given in (11) for two characteristics satisfying the assumptions of Proposition 3.1 and for the case when the covariance matrix of rebalancing trades is symmetric with respect to the two characteristics; that is, the case when $\sigma_1 = \sigma_2$. Panel (a) shows the price-impact diversification ratio on the vertical axis as a function of the correlation between the rebalancing trades of the two characteristics on the horizontal axis. The two curves correspond to the cases with square-root price-impact function (subquadratic price-impact costs, $\alpha = 0.5$) and linear price-impact function (quadratic price-impact costs, $\alpha = 1$). Panel (b) shows the price-impact diversification ratio on the vertical axis as a function of the number of characteristics combined for the case with $\rho = 0$ and square-root price-impact function ($\alpha = 0.5$).

(a) Effect of rebalancing trade correlation for two characteristics



(b) Effect of number of characteristics for zero rebalancing-trade correlation



The explanation for these theoretical findings is that the subquadratic price-impact cost function ($\alpha = 0.5$) assigns a lower cost to large trades than the quadratic function ($\alpha = 1$). To see this, consider a simple example with two characteristics whose rebalancing trades in the nth stock, $\tilde{x}_{1tn}$ and $\tilde{x}_{2tn}$, are independently and identically distributed with equal probability to take a value of -1 or $+1$. Moreover, let the price-impact parameter for the nth stock be constant and equal to one, $\lambda_{tn} = 1$.[10] Then, for the case with quadratic price-impact costs, the expected cost of rebalancing each of the characteristics in isolation is equal to one:

$$E\left[|\tilde{x}_{ktn}|^2\right] = \frac{1}{2}|-1|^2 + \frac{1}{2}|+1|^2 = 1 \quad \text{for} \quad k = 1, 2.$$

To calculate the price-impact cost of trading the two characteristics in combination, we consider four equally likely outcomes depending on the values of $\tilde{x}_{1tn}$ and $\tilde{x}_{2tn}$:

$$(\tilde{x}_{1tn}, \tilde{x}_{2tn}) = \begin{cases} (-1, -1), \\ (-1, +1), \\ (+1, -1), \\ (+1, +1). \end{cases}$$

Thus, the expected cost of trading the two characteristics in combination is equal to two:

$$\begin{aligned} E\left[|\tilde{x}_{1tn} + \tilde{x}_{2tn}|^2\right] &= \frac{1}{4} \times |-1-1|^2 + \frac{1}{4} \times |-1+1|^2 + \frac{1}{4} \times |+1-1|^2 + \frac{1}{4} \times |+1+1|^2 \\ &= \frac{1}{4} \times |-2|^2 + \frac{1}{4} \times |0|^2 + \frac{1}{4} \times |0|^2 + \frac{1}{4} \times |+2|^2 = 2. \end{aligned}$$

Thus, from (11) we have that

$$\text{price-impact diversification ratio} = \frac{E\left[|\tilde{x}_{1tn} + \tilde{x}_{2tn}|^2\right]}{E\left[|\tilde{x}_{1tn}|^2\right] + E\left[|\tilde{x}_{2tn}|^2\right]} = \frac{2}{1+1} = 1.$$

The price-impact diversification ratio is one for the case with quadratic costs because even though the price-impact cost is zero for the two outcomes where the trades of the two characteristics net out, $(-1, +1)$ and $(+1, -1)$, the *quadratic* price-impact costs of trading

[10] We assume that the absolute value of the rebalancing trades and the price-impact parameter are equal to one without loss of generality as the price-impact diversification ratio is invariant to multiplying the rebalancing trades of all characteristics or the price-impact parameter by a constant because the absolute-value and power functions are homogeneous.

the characteristics in combination are large (equal to four) for the two outcomes where the rebalancing trades of the two characteristics are in the same direction, $(-1,-1)$ and $(+1,+1)$. In other words, the increase in price-impact cost from trading characteristics in combination for the two outcomes where the trades of the two characteristics are in the same direction exactly compensates for the reduction for the two outcomes where they net out. This is because the quadratic price-impact costs are disproportionately high for large trades.

For the square-root price-impact function, that is, subquadratic price-impact costs, the expected cost of rebalancing each of the characteristics in isolation is also equal to one:

$$E\left[|\tilde{x}_{ktn}|^{1.5}\right] = \frac{1}{2} \times |-1|^{1.5} + \frac{1}{2} \times |+1|^{1.5} = 1 \quad \text{for} \quad k = 1, 2.$$

In contrast, the expected cost of trading the two characteristics in combination is only $\sqrt{2}$:

$$\begin{aligned} E\left[|\tilde{x}_{1tn} + \tilde{x}_{2tn}|^{1.5}\right] &= \frac{1}{4} \times |-1-1|^{1.5} + \frac{1}{4} \times |-1+1|^{1.5} + \frac{1}{4} \times |+1-1|^{1.5} + \frac{1}{4} \times |+1+1|^{1.5} \\ &= \frac{1}{4} \times |-2|^{1.5} + \frac{1}{4} \times |0|^{1.5} + \frac{1}{4} \times |0|^{1.5} + \frac{1}{4} \times |+2|^{1.5} \\ &= \frac{2 \times 2^{1.5}}{4} = \sqrt{2}. \end{aligned}$$

Thus, we have that for the square-root price-impact function the

$$\text{price-impact diversification ratio} = \frac{E\left[|\tilde{x}_{1tn} + \tilde{x}_{2tn}|^{1.5}\right]}{E\left[|\tilde{x}_{1tn}|^{1.5}\right] + E\left[|\tilde{x}_{2tn}|^{1.5}\right]} = \frac{\sqrt{2}}{1+1} = \frac{1}{\sqrt{2}} < 1.$$

The price-impact diversification ratio is smaller than one for the case with square-root price impact because, in addition to having zero price-impact cost for the two outcomes where the rebalancing trades of the two characteristics cancel out, $(-1,+1)$ and $(+1,-1)$, the *subquadratic* price-impact costs of trading the characteristics in combination are only $2^{1.5}$ for the two outcomes where the rebalancing trades of the two characteristics are in the same direction, $(-1,-1)$ and $(+1,+1)$, compared to 2^2 for the case with quadratic costs.

This example illustrates why, with a square-root price impact, combining characteristics leads to a reduction of price-impact costs even when the rebalancing trades of the two

Table 1: List of characteristics considered

This table lists the 18 characteristics we consider, which include the traditional characteristics size, value, and momentum, plus the 15 characteristics that DeMiguel et al. (2020) find to be significant, ordered alphabetically by acronym. The first column gives the number of the characteristic, the second column gives the characteristic's definition, the third column gives the acronym, and the fourth and fifth columns give the authors who analyzed them, and the date and journal of publication. Our definitions and acronyms match those in Green, Hand, and Zhang (2017).

#	Characteristic and definition	Acronym	Author(s)	Date, Journal
1	Asset growth: Annual percent change in total assets	agr	Cooper, Gulen & Schill	2008, JF
2	Beta: Estimated market beta from weekly returns and equal weighted market returns for 3 years ending month $t-1$ with at least 52 weeks of returns	beta	Fama & MacBeth	1973, JPE
3	Book to market: Book value of equity divided by end of fiscal-year market capitalization	bm	Rosenberg, Reid & Lanstein	1985, JPM
4	Industry adjusted book to market: Industry adjusted book-to-market ratio	bm_ia	Asness, Porter & Stevens	2000, WP
5	Industry adjusted change in asset turnover: 2-digit SIC fiscal-year mean adjusted change in sales divided by average total assets	chatoia	Soliman	2008, TAR
6	Change in tax expense: Percent change in total taxes from quarter $t-4$ to t	chtx	Thomas & Zhang	2011 JAR
7	Gross profitability: Revenues minus cost of goods sold divided by lagged total assets	gma	Novy-Marx	2013 JFE
8	Industry sales concentration: Sum of squared percent of sales in industry for each company	herf	Hou & Robinson	2006, JF
9	12-month momentum: 11-month cumulative returns ending one month before month-end	mom12m	Jegadeesh	1990, JF
10	1-month momentum: 1-month cumulative return	mom1m	Jegadeesh	1990, JF
11	Market capitalization: Natural log of market capitalization at end of month $t-1$	mve	Banz	1981, JFE
12	Δ% gross margin - Δ% sales: Percent change in gross margin minus percent change in sales	pchgm_pchsale	Abarbanell & Bushee	1998, TAR
13	Financial-statements score: Sum of 9 indicator variables to form fundamental health score	ps	Piotroski	2000, JAR
14	R&D to market cap: R&D expense divided by end-of-fiscal-year market capitalization	rd_mve	Guo, Lev & Shi	2006, JBFA
15	Return volatility: Standard deviation of daily returns from month $t-1$	retvol	Ang, Hodrick, Xing & Zhang	2006, JF
16	Volatility of share turnover: Monthly standard deviation of daily share turnover	std_turn	Chordia, Subrahmanyan & Anshuman	2001, JFE
17	Unexpected quarterly earnings: Unexpected quarterly earnings divided by fiscal-quarter-end market cap. Unexpected earnings is I/B/E/S actual earnings minus median forecasted earnings if available, else it is the seasonally differenced quarterly earnings before extraordinary items from Compustat quarterly file	sue	Rendelman, Jones & Latane	1982, JFE
18	Zero trading days: Turnover weighted number of zero trading days for most recent month	zerotrade	Liu	2006, JFE

characteristics are uncorrelated or moderately positively correlated. In the next section, we examine the magnitude of this effect empirically.

3.2 Empirical results

To evaluate empirically the trading diversification benefits from combining characteristics, we require a historical sample of the rebalancing-trade vectors, $\tilde{x}_{kt}$, the characteristic portfolio returns, $\tilde{x}_{kt}^\top r_{t+1}$, and an estimate of the price-impact cost for each stock.

To obtain a historical sample for the rebalancing-trade vectors and characteristic-portfolio returns, we compile stock-return data as well as data for the 18 characteristics listed in Table 1, which include the traditional characteristics size, value, and momentum plus the 15 characteristics that DeMiguel et al. (2020) find jointly significant for explaining the cross section of stock returns. We combine U.S. stock-market information from CRSP, Compustat, and I/B/E/S from January 1980 to December 2017. Our database contains every firm traded on the NYSE, AMEX, and NASDAQ exchanges. We then remove firms with negative book-to-market ratios. As in Brandt et al. (2009), we also remove firms below the 20th percentile of market capitalization because these are very small firms that are difficult to trade. We form value-weighted long-short portfolios for each characteristic by going long on stocks with values of the characteristic above the 30th percentile and going short stocks with values of the characteristic below the 70th percentile. We standardize the value-weights so that both the positive and negative weights sum to one for each characteristic. Then, we use (6) to compute the monthly rebalancing-trade vectors and a single-characteristic version of (2) to compute the returns of each characteristic portfolio.

To estimate the price-impact cost for the nth stock, we rely on the results by Frazzini et al. (2018), who use a trade-execution dataset from a large institutional money manager covering a 19-year period to estimate the following panel regression for the price impact of a trade on the nth stock

$$\mathrm{PI}_{tn} = a_{tn} + b\ \mathrm{vr}_{tn} + c\ \mathrm{sign}(\mathrm{vr}_{tn})\sqrt{|\mathrm{vr}_{tn}|}, \tag{14}$$

where $\mathrm{vr}_{tn} = 100 \times \Delta w_{tn}/\mathrm{dtv}_{tn}$ is the signed dollar value of a trade Δw_{tn} as a percentage of the stock's average daily dollar volume dtv_{tn}. The second and third terms on the right-hand side of (14) represent linear and square-root price impact, respectively. The first term a_{tn} captures the effect of explanatory variables that do not depend on the trade size such as a time trend, the log market capitalization and idiosyncratic volatility of the stock, and the monthly variance of the CRSP-value weighted index. The panel regression in (14) is a generalization of the price-impact function given in Equation (3) because it allows for additional explanatory variables collected in a_{tn}, and a term linear in vr_{tn} in addition to a square-root term. Frazzini et al. (2018) find that the coefficient c in (14) is highly statistically significant, whereas the coefficient b is not significant, consistent with the findings of Torre and Ferrari (1997), Grinold and Kahn (2000), Almgren et al. (2005), and Ratcliffe et al. (2017). We rely on the estimates of a_{tn}, b, and c reported in Column (9) of Table VII in Frazzini et al. (2018), to predict the price-impact cost of a trade in the nth stock.

Given the historical sample of rebalancing-trade vectors and the stock price-impact model of Frazzini et al. (2018), Table 2 compares the capacity, optimal investment, and optimal annual profit associated with exploiting the 18 characteristics when considered in isolation and in combination. We obtain the optimal investment and profit by numerically solving the parametric portfolio problem (10) for each of the characteristics in isolation and for the 18 in combination, with the price-impact cost PIC_t evaluated using the model of Frazzini et al. (2018) in Equation (14). The investment is given by the optimal value of θ and the *annual* profit is 12 times the optimal objective of problem (10). We obtain the capacity of each characteristic in isolation by computing the maximum investment that can be allocated to each characteristic before price-impact costs erode any profits. We obtain the capacity of the 18 characteristics in combination by scaling up the optimal parametric portfolio θ for the case where the 18 characteristics are exploited in combination until price-impact costs erode any profits. Like Korajczyk and Sadka (2004) and Novy-Marx and Velikov (2016), we express all quantities in terms of market capitalization at the end of our sample (December 2017).

The estimates of capacity in Table 2 are consistent with those in the existing literature; for instance, the total capacity aggregated across the 18 characteristics when considered in isolation is around 200 billion dollars. More importantly, the table shows that trading diversification has a first-order effect on capacity, investment, and profits. In particular, we observe that, when the characteristics are exploited in combination, total capacity and total optimal investment increase by around 50% and total profits increase by around 25%.

Table 2: Capacity, investment, and profit in isolation and combination

This table reports the capacity, investment, and profit of each characteristic when considered in isolation and in combination. For each characteristic, the first column reports its acronym and the remaining columns report its capacity, optimal investment, and optimal profit when considered in isolation and in combination, as well as the percentage increase in these quantities when the characteristic is considered in combination instead of in isolation. We obtain the optimal investment and profit by numerically solving problem (10) for each of the characteristics in isolation and for the 18 in combination, with the price-impact cost PIC_t evaluated using the model of Frazzini et al. (2018) in Equation (14). The investment is given by the optimal value of θ and the annual profit is 12 times the optimal objective of problem (10). We obtain the capacity of each characteristic in isolation by computing the maximum investment that can be allocated to each characteristic before price-impact costs erode any profits. We obtain the capacity of the 18 characteristics in combination by scaling up the optimal parametric portfolio θ for the case where the 18 characteristics are exploited in combination until price-impact costs erode any profits. Like Korajczyk and Sadka (2004) and Novy-Marx and Velikov (2016), we express all quantities in terms of market capitalization at the end of our sample (December 2017).

	Capacity			Investment			Profit		
Characteristic	Isol. ($bill.)	Comb. ($bill.)	Incr. (%)	Isol. ($bill.)	Comb. ($bill.)	Incr. (%)	Isol. ($mill.)	Comb. ($mill.)	Incr. (%)
gma	76.410	101.973	33	36.665	49.690	36	208.80	308.65	48
rd_mve	51.836	51.001	−2	25.918	24.852	−4	686.08	681.92	−1
bm	34.617	42.938	24	16.474	20.923	27	163.61	215.18	32
herf	2.225	27.024	1115	1.025	13.168	1184	0.44	13.03	2878
agr	14.586	22.358	53	6.863	10.895	59	65.33	119.91	84
chatoia	4.117	9.257	125	1.925	4.511	134	9.91	31.08	214
ps	5.478	7.285	33	2.542	3.550	40	15.54	25.91	67
bm_ia	0.000	6.296	-	0.000	3.068	-	0.00	1.32	-
pchgm_pchsale	2.489	4.206	69	1.144	2.050	79	4.29	10.63	148
beta	0.331	3.343	909	0.157	1.629	937	0.13	3.70	2847
mom12m	1.319	3.236	145	0.608	1.577	159	5.48	18.41	236
sue	1.083	2.198	103	0.497	1.071	116	3.73	11.16	199
chtx	0.821	2.013	145	0.379	0.981	159	1.77	7.82	341
retvol	0.246	1.554	531	0.111	0.757	580	0.78	9.71	1140
std_turn	0.000	1.091	-	0.000	0.532	-	0.00	2.18	-
mom1m	0.000	0.522	-	0.000	0.254	-	0.00	3.18	-
zerotrade	0.006	0.365	5974	0.003	0.178	6184	0.00	1.30	38400
mve	0.000	0.198	-	0.000	0.096	-	0.00	0.06	-
Total	195.564	286.859	47	94.312	139.782	48	1165.89	1465.15	26

Figure 3 reports the empirical price-impact diversification ratio as a function of the number of characteristics considered, when we invest in each characteristic the amount that is optimal when all 18 characteristics are considered in combination; that is, the amount in the sixth column of Table 2.[11] For each number of characteristics $K = 1, 2, \ldots, 18$ (depicted on the horizontal axis), we consider all combinations of the 18 characteristics taken in groups of K and report the 10th percentile, mean, and 90th percentile of price-impact diversification ratio across the combinations. Figure 3 shows that the benefits from trading diversification increase substantially with the number of characteristics, and the price-impact diversification ratio is around 83% for the case where all 18 characteristics are considered; that is, there is reduction of around 17% in price-impact costs when combining all characteristics using their optimal investment weights compared to considering them in isolation.[12] Moreover, Figure 3 shows that the results in Table 2 are robust. To see this, note that we do not re-optimize the weights of the characteristics when considering only $K < 18$, but rather use the weights that are optimal when combining all 18 characteristics, yet there are substantial benefits in terms of price-impact cost from combining characteristics.

In this section, we have shown empirically that combining characteristics substantially reduces the price impact of exploiting them. In the remainder of this manuscript, we study how price-impact diversification affects the strategic interaction between investors and the resulting effects on investment positions and profits. For simplicity, we consider two groups of investors, each group exploiting a different characteristic.

4 Game-theoretic model of strategic competition

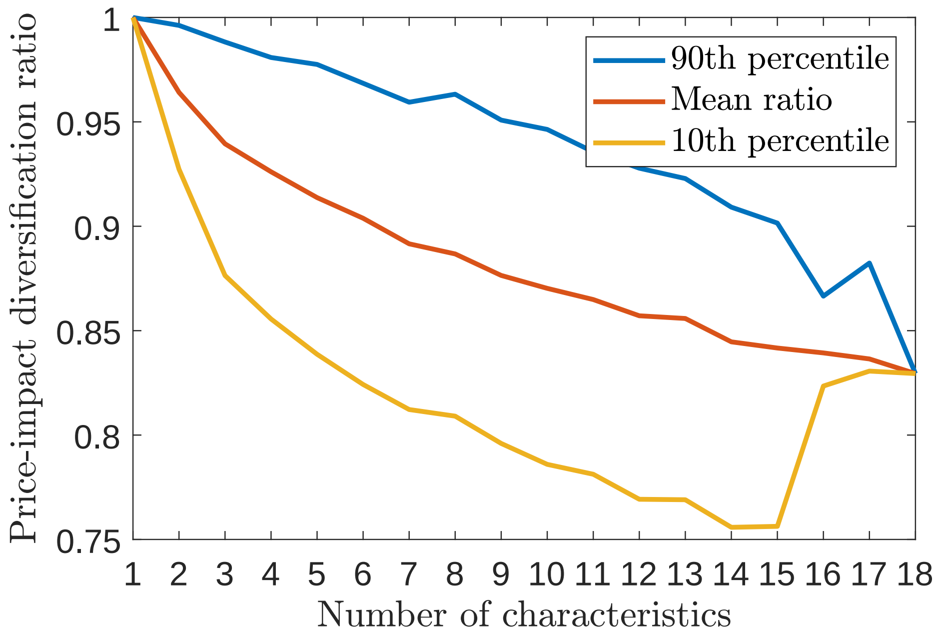
We now extend the parametric-portfolio framework introduced in Section 2 to develop a game-theoretic model of competition where I_1 investors exploit the first characteristic and

[11] We have re-produced Figure 3 for the case where the characteristics are equally weighted as in Equation (11) and the results are very similar with a price-impact cost reduction of around 19% from combining 18 characteristics, compared to exploiting them in isolation.

[12] Although substantial, this price-impact cost reduction is smaller than that predicted by Proposition 3.1 because the multivariate distribution of the *empirical* rebalancing trades is neither symmetric nor normally distributed.

Figure 3: Price-impact diversification ratio versus number of characteristics

This figure depicts the empirical price-impact diversification ratio as a function of the number of characteristics combined, when we invest in each characteristic the amount that is optimal when all 18 characteristics are considered in combination. For each number of characteristics K (depicted on the horizontal axis), we consider all combinations of the 18 characteristics taken in groups of K and report the 10th percentile, mean, and 90th percentile of price-impact diversification ratio across the combinations.



I_2 the second. Section 5 uses this model to study how price-impact diversification affects the strategic interactions between investors and Section 6 takes the model to the data.

The portfolio of the ith investor exploiting the kth characteristic at time t is

$$w_{kit}(\theta_{ki}) = x_{kt}\theta_{ki},$$

where $\theta_{ki} \in \mathbb{R}$ is the investment position of the ith investor exploiting the kth characteristic. The portfolio return of the ith investor exploiting the kth characteristic at time $t+1$ is

$$r_{ki,t+1} = \theta_{ki} x_{kt}^{\top} r_{t+1}. \tag{15}$$

For analytical tractability, in this section we assume that trading has a linear impact on prices; that is, $\alpha = 1$, an assumption that we relax in our empirical work in Section 6. The following lemma shows that, under the assumption of a linear price-impact function, the price-impact cost of the investors at time t is a quadratic function of their investment positions.

Lemma 4.1 *Assume the aggregate amount of trading on a stock has a linear impact on its price ($\alpha = 1$). Then, the price-impact cost of the ith investor exploiting the first and second characteristics at time t can be written as*

$$PIC_{1it} = \theta_{1i}\lambda_{1t}(\theta_{1i} + \theta_{1,-i}) + \theta_{1i}\lambda_{12t}\sum_{j=1}^{I_2}\theta_{2j} \quad \text{and} \tag{16}$$

$$PIC_{2it} = \theta_{2i}\lambda_{2t}(\theta_{2i} + \theta_{2,-i}) + \theta_{2i}\lambda_{12t}\sum_{j=1}^{I_1}\theta_{1j}, \quad \text{respectively,} \tag{17}$$

where $\theta_{k,-i} = \sum_{j\neq i}\theta_{kj}$ is the aggregate investment position of investors in the kth characteristic other than the ith investor and $\lambda_{kt} = \tilde{x}_{kt}^{\top}\Lambda_t\tilde{x}_{kt}$ and $\lambda_{12t} = \tilde{x}_{1t}^{\top}\Lambda_t\tilde{x}_{2t}$ are the price-impact parameters for the kth characteristic and the interaction between the two characteristics at time t, respectively, where $\tilde{x}_{kt}$ is the rebalancing-trade vector for the kth characteristic at time t defined in (8)*.*

Lemma 4.1 shows that, for the case with linear price-impact function, the trading costs for the two investors can be conveniently decomposed into three separate terms. The terms associated with λ_{1t} and λ_{2t} measure the price-impact cost associated with exploiting in isolation the first and second characteristics, respectively. The parameter λ_{12t} measures the *interaction* between the rebalancing trades for the two characteristics. For $\lambda_{12t} = 0$, the price-impact costs of exploiting the two characteristics are independent, for $\lambda_{12t} < 0$ (> 0) there is a positive (negative) externality between the two groups of investors because trading in one characteristic decreases (increases) the price-impact cost of trading the other.

4.1 Decentralized setting

In the decentralized setting, we consider a game where the two groups of investors make decisions *simultaneously*; however, in unreported results we observe that our findings are robust to considering the case where investors in one of the groups act as Stackelberg leaders.[13]

[13]Also, as mentioned in Section 2, although for simplicity of exposition we assume investors are risk neutral, Appendix IA.1.1 shows that our results are robust to considering risk-averse investors.

The ith investor in the kth characteristic chooses her investment position θ_{ki} to optimize the unconditional expectation of the difference between her price-impact cost and portfolio return

$$\min_{\theta_{ki}} \quad E[PIC_{kit} - r_{ki,t+1}]. \tag{18}$$

Then, using (15) and (16), the decision problem of the ith investor in the first and second characteristics can be rewritten as

$$\min_{\theta_{1i}} \quad E[\theta_{1i}\lambda_{1t}(\theta_{1i} + \theta_{1,-i}) + \theta_{1i}\lambda_{12t}\sum_{j=1}^{I_2}\theta_{2j} - \theta_{1i}x_{1t}^{\top}r_{t+1}], \quad \text{and} \tag{19}$$

$$\min_{\theta_{2i}} \quad E[\theta_{2i}\lambda_{2t}(\theta_{2i} + \theta_{2,-i}) + \theta_{2i}\lambda_{12t}\sum_{j=1}^{I_1}\theta_{1j} - \theta_{2i}x_{2t}^{\top}r_{t+1}], \quad \text{respectively.} \tag{20}$$

These decision problems can be rewritten as

$$\min_{\theta_{1i}} \quad \theta_{1i}\lambda_1(\theta_{1i} + \theta_{1,-i}) + \theta_{1i}\lambda_{12}\sum_{j=1}^{I_2}\theta_{2j} - \theta_{1i}\mu_1 \quad \text{and} \tag{21}$$

$$\min_{\theta_{2i}} \quad \theta_{2i}\lambda_2(\theta_{2i} + \theta_{2,-i}) + \theta_{2i}\lambda_{12}\sum_{j=1}^{I_1}\theta_{1j} - \theta_{2i}\mu_2, \quad \text{respectively,} \tag{22}$$

where $\lambda_k = E[\lambda_{kt}]$ is the price-impact parameter for the kth characteristic, $\lambda_{12} = E[\lambda_{12t}]$ is the price-impact parameter for the interaction between the two characteristics, and $\mu_k = E[x_{kt}^{\top}r_{t+1}]$ is the average return of the kth characteristic portfolio. Note that although μ_1 and μ_2 are exogenous to our model, the average characteristic return *net of price-impact costs*, $\bar{\mu}_1$ and $\bar{\mu}_2$, are determined endogenously as a function of the investment positions. For instance, $\bar{\mu}_1 = \mu_1 - \lambda_1(\theta_{1i} + \theta_{1,-i}) - \lambda_{12}\sum_{j=1}^{I_2}\theta_{2j}$.

4.2 Centralized setting

To understand the impact of competition between the two groups of investors, we also consider a centralized setting in which a single investor exploits both characteristics. For the case with linear price impact, the decision problem in the centralized setting is:

$$\min_{\theta_{1c},\theta_{2c}} \quad \theta_{1c}\lambda_1\theta_{1c} + 2\theta_{1c}\lambda_{12}\theta_{2c} + \theta_{2c}\lambda_2\theta_{2c} - \theta_{1c}\mu_1 - \theta_{2c}\mu_2, \tag{23}$$

where the subscript "c" denotes the optimal quantities for the centralized market. Note that because the objective function in the centralized setting is to maximize total profits, the total profit in the centralized setting is an upper bound for that in the decentralized setting.

4.3 Equilibrium

We now characterize the unique equilibrium in closed form for both the decentralized and centralized settings. We start with some assumptions that rule out unrealistic cases.

Assumption 4.1 *The joint probability distribution of the two characteristic rebalancing-trade vectors, $\tilde{x}_{1t}$ and $\tilde{x}_{2t}$, is such that the following events have strictly positive probability:*

1. *The rebalancing-trade vector of the first characteristic is nonzero; that is, $\tilde{x}_{1t} \neq 0$.*
2. *The rebalancing-trade vector of the second characteristic is nonzero; that is, $\tilde{x}_{2t} \neq 0$.*
3. *The rebalancing-trade vectors of the two characteristics are not equal, up to a change of scale; that is, there does not exist $a \in \mathbb{R}$ such that $\tilde{x}_{1t} = a\tilde{x}_{2t}$.*

Assumptions 4.1(1–2) rule out the case in which exploiting the characteristics does not require any rebalancing trades. Assumption 4.1(3) rules out the case in which the two characteristics require rebalancing trades that are identical, up to a change of scale.

Under Assumption 4.1, we show using the triangular inequality that the absolute value of the price-impact parameter for the interaction between the two characteristics is bounded above.

Lemma 4.2 *Let Assumption 4.1 hold. Then, $\lambda_1, \lambda_2 > 0$ and the absolute value of the price-impact parameter for the interaction between the two characteristics is bounded above by $\bar{\lambda}_{12} \equiv \sqrt{\lambda_1 \lambda_2}$; that is, $|\lambda_{12}| < \bar{\lambda}_{12}$.*

The following proposition provides closed-form expressions for the equilibrium quantities in the *decentralized* setting, denoted by the subscript "d."

Proposition 4.1 *Let Assumption 4.1 hold, then in the decentralized setting:*

1. *There exists a unique Nash equilibrium.*

2. *The equilibrium is symmetric with respect to the I_1 investors exploiting the first characteristic and with respect to the I_2 investors exploiting the second.*

3. *The investment positions of the ith investor exploiting the first characteristic and the ith investor exploiting the second characteristic are*

$$\theta_{1id} = \frac{(I_2+1)\,\lambda_2\mu_1 - I_2\lambda_{12}\mu_2}{(I_1+1)(I_2+1)\lambda_1\lambda_2 - I_1I_2\lambda_{12}^2} \quad \text{and} \tag{24}$$

$$\theta_{2id} = \frac{(I_1+1)\lambda_1\mu_2 - I_1\lambda_{12}\mu_1}{(I_1+1)(I_2+1)\lambda_1\lambda_2 - I_1I_2\lambda_{12}^2}, \quad \text{respectively.} \tag{25}$$

4. *The profits of the ith investor exploiting the kth characteristic is*

$$\pi_{kid} = \lambda_k\theta_{kid}^2. \tag{26}$$

The following proposition gives the optimal investments and profit in the centralized setting, denoted by the subscript "c".

Proposition 4.2 *Let Assumption 4.1 hold, then in the centralized setting:*

1. *There exists a unique minimizer to the centralized decision problem.*

2. *The optimal investment positions are*

$$\theta_{1c} = \frac{\lambda_2\mu_1 - \lambda_{12}\mu_2}{2(\lambda_1\lambda_2 - \lambda_{12}^2)}, \tag{27}$$

$$\theta_{2c} = \frac{\lambda_1\mu_2 - \lambda_{12}\mu_1}{2(\lambda_1\lambda_2 - \lambda_{12}^2)}. \tag{28}$$

3. *The profits from the first and second characteristics are*

$$\pi_{1c} = \frac{\frac{1}{2}\lambda_2\mu_1^2 - \frac{1}{2}\lambda_{12}\mu_1\mu_2}{2(\lambda_1\lambda_2 - \lambda_{12}^2)}, \tag{29}$$

$$\pi_{2c} = \frac{\frac{1}{2}\lambda_1\mu_2^2 - \frac{1}{2}\lambda_{12}\mu_1\mu_2}{2(\lambda_1\lambda_2 - \lambda_{12}^2)}. \tag{30}$$

Propositions 4.1 and 4.2 characterize the equilibrium in closed form for the general case where both characteristics may have a nonzero mean return. For simplicity of exposition, Section 5 studies the case where the average return of the second characteristic is zero ($\mu_2 = 0$). We refer to this case as the case with *pure liquidity-provision motive* because in this case the only motive to trade the second characteristic is to receive compensation for providing liquidity for the trades of the first characteristic. The empirical analysis in Section 6 shows that our findings are robust to the general case with $\mu_2 \neq 0$. The following assumption describes the case with pure liquidity-provision motive.

Assumption 4.2 *The mean return of the first characteristic is strictly positive, $\mu_1 > 0$, the mean return of the second characteristic is zero, $\mu_2 = 0$, and the price-impact parameter for the interaction between the two characteristics is positive, $\lambda_{12} > 0$.*

Note that the assumption that characteristic mean returns are nonnegative is without loss of generality because the case where a characteristic has a negative mean return can be transformed into a case with positive mean return by changing the sign of the characteristic. Also, we exclude the trivial case with $\lambda_{12} = 0$, in which the decisions of the two groups of investors are independent. Finally, it is straightforward to show that for the case with $\mu_2 = 0$, the equilibrium quantities are the same for the cases with $\lambda_{12} > 0$ and $\lambda_{12} < 0$, up to a change of sign; therefore, the assumption that $\lambda_{12} > 0$ is without loss of generality.

5 Discussion of equilibrium

Our discussion of the equilibrium of the game-theoretic model parallels our discussion of Figure 1 in the introduction. To study the effect of crowding, we start by considering the case where there are only investors exploiting the first characteristic ($I_1 \geq 1$, $I_2 = 0$). Then, to characterize how trading diversification and competition among investors exploiting the second characteristic alleviate crowding in the first characteristic, we consider the case where there are investors exploiting both characteristics ($I_1 \geq 1$ and $I_2 \geq 1$) . Finally, we consider the centralized setting where a single investor exploits both characteristics.

5.1 Crowding in factor investing

To set the stage for our main insight about trading diversification, we begin by considering the case where there are only investors exploiting the first characteristic ($I_1 \geq 1$ and $I_2 = 0$). We identify the *capacity* of the first characteristic, defined as the aggregate investment position for which aggregate profits become zero. From Equation (21), we have that the aggregate profits for the case where there are no investors exploiting the second characteristic are $\pi_1 = \theta_1\mu_1 - \theta_1\lambda_1\theta_1$, where the aggregate investment position in the first characteristic is $\theta_1 = \sum_{i=1}^{I_1} \theta_{1i}$. Thus, the capacity of the first characteristic is $C(I_2 = 0) = \mu_1/\lambda_1$.

In the following proposition, we characterize analytically the equilibrium for the case where we only have investors exploiting the first characteristic and there are no investors exploiting the second characteristic.

Proposition 5.1 *Let Assumption 4.1 hold and consider the case where there are only investors exploiting the first characteristic ($I_1 \geq 1$ and $I_2 = 0$), then there exists a unique Nash equilibrium, which is symmetric across the I_1 investors. Moreover, the aggregate investment position of the investors in the first characteristic is*

$$\theta_{1d} = I_1\theta_{1id} = \frac{I_1}{I_1 + 1}\frac{\mu_1}{\lambda_1} \tag{31}$$

and their aggregate profits are

$$\pi_{1d} = I_1\pi_{1id} = I_1\lambda_1\theta_{1id}^2 = \frac{I_1}{(I_1 + 1)^2}\frac{\mu_1^2}{\lambda_1}. \tag{32}$$

Furthermore, the following monotonicity properties hold:

1. *The aggregate investment position $\theta_{1d} = I_1\theta_{1id}$ is increasing in I_1 and converges to the strategy's capacity μ_1/λ_1 as $I_1 \to \infty$.*

2. *The aggregate profits $\pi_{1d} = I_1\pi_{1id}$ are decreasing in I_1 and converge to zero as $I_1 \to \infty$.*

The intuition underlying Proposition 5.1 is as follows. A single investor maximizes her profits by investing *half* of the capacity, $\theta_{1d} = \frac{1}{2}\frac{\mu_1}{\lambda_1} = \frac{C(I_2=0)}{2}$. Note that this is the first-best allocation that maximizes aggregate profits in the absence of a second characteristic because the single investor acts as a monopolist exploiting the first characteristic. However, when there are multiple investors competing to exploit the first characteristic, there is a negative externality among them because they do not internalize in their objective function how their investment decisions affect each other. Consequently, as the number of investors I_1 increases, their aggregate investment position increases and their aggregate profit decreases because the externality among them worsens. In the limit, as the number of investors goes to infinity, the externality pushes them to overinvest to the point where price-impact costs completely erode any profits from trading the first characteristic.[14] Thus, Proposition 5.1 establishes the base-case result that *competition among investors exploiting the same characteristic erodes their profits because of crowding.*

5.2 Trading diversification

To study the effect of trading diversification, we now consider the case where there may also be investors exploiting the second characteristic ($I_2 \geq 0$). We first characterize how trading diversification and competition among investors exploiting the second characteristic increase the capacity of the first characteristic.

Proposition 5.2 *Let Assumptions 4.1 and 4.2 hold and $\lambda_{12} > 0$, then the capacity of the first characteristic for any $I_2 \geq 0$ is*

$$C(I_2) = \frac{\mu_1}{\lambda_1 - \frac{I_2}{I_2+1} \frac{\lambda_{12}^2}{\lambda_2}}.$$

Moreover, $C(I_2)$ is monotonically increasing in I_2 for any $I_2 \geq 0$.

We then characterize how the equilibrium aggregate investment position and profits for the first characteristic increase with the number of investors exploiting the second characteristic, I_2.

[14]This result parallels the classic result of competition in quantities first studied by Cournot (1838).

Proposition 5.3 *Let Assumptions 4.1 and 4.2 hold and $I_1 < \infty$, then the equilibrium quantities in the decentralized setting given in Proposition 4.1 satisfy the following conditions with respect to the number of investors exploiting the second characteristic, $I_2 \geq 0$:*

1. *The aggregate investment position in the first characteristic $\theta_{1d} = I_1\theta_{1id}$ is strictly positive and increasing in I_2.*
2. *The aggregate profit from the first characteristic $\pi_{1d} = I_1\pi_{1id}$ is strictly positive and increasing in I_2.*
3. *For $I_2 \geq 1$, the investment position in the second characteristic $\theta_{2d} = I_2\theta_{2id}$ is strictly negative and decreasing in I_2; that is, it is increasing in absolute value.*
4. *For $I_2 \geq 1$, the aggregate profit from the second characteristic $\pi_{2d} = I_2\pi_{2id}$ is strictly positive and decreasing in I_2 provided that $\frac{2-(I_2+1)}{I_2} > \frac{I_1}{I_1+1}$, and converges to zero as $I_2 \to \infty$.*

We now discuss the intuition underlying Propositions 5.2 and 5.3, which are illustrated in Panel (b) of Figure 1. First, comparing the case where there is no investor ($I_2 = 0$) to that where there is a single investor ($I_2 = 1$) exploiting the second characteristic, Propositions 5.2 and 5.3 show that trading diversification increases the capacity as well as the equilibrium aggregate investment position and profits of the first characteristic. *Thus, trading diversification alleviates crowding in the first characteristic.* Second, an increase in competition among investors exploiting the second characteristic, measured by an increase in I_2, further increases the capacity as well as the equilibrium aggregate investment position and profits for the first characteristic. To understand this result, note that there is a negative externality among the investors exploiting the second characteristic because they do not internalize in their objective function the effect of their investment decisions on each other. This externality worsens as I_2 increases and leads them to increase their aggregate investment position, which reduces their profits, but increases aggregate profits from the first characteristic. *Thus, competition among investors exploiting the second characteristic further alleviates crowding in the first characteristic.*

Section 5.1 showed that in the absence of investors exploiting the second characteristic, increased competition among investors exploiting the first characteristic leads to an increase in their aggregate investment position and a decrease in their aggregate profits. The following proposition shows that this monotonicity result with respect to I_1 holds also when there are $I_2 \geq 1$ investors exploiting the second characteristic.

Proposition 5.4 *Let Assumptions 4.1 and 4.2 hold and $I_2 < \infty$, then the decentralized equilibrium quantities in Proposition 4.1 satisfy the following conditions with respect to I_1:*

1. *The aggregate investment position in the first characteristic $\theta_{1d} = I_1\theta_{1id}$ is increasing in I_1.*
2. *The aggregate investment position in the second characteristic $\theta_{2d} = I_2\theta_{2id}$ is strictly negative and decreasing in I_1 for $\lambda_{12} > 0$; that is, it is increasing in absolute value.*
3. *The aggregate profits from the first characteristic $\pi_{1d} = I_1\pi_{1id}$ are decreasing in I_1 for I_1 such that $(I_1 - 1)/I_1 \geq I_2/(I_2 + 1)$ and converge to zero as $I_1 \to \infty$.*
4. *The aggregate profits from the second characteristic $\pi_{2d} = I_2\pi_{2id}$ are strictly positive and increasing in I_1 for $\lambda_{12} > 0$.*

Proposition 5.4 shows that, even when there are investors exploiting the second characteristic, competition among investors exploiting the first characteristic leads to a reduction in their profits. However, competition among investors exploiting the first characteristic also leads investors exploiting the second characteristic to increase their investment positions. This is because the increased investment position in the first characteristic increases the rents from exploiting the second characteristic because of the positive externality between investors exploiting the two characteristics. This increases the market power of the investors exploiting the second characteristic, who strategically increase their investment positions and earn higher profits. Thus, although competition among investors exploiting the first characteristic erodes their profits because of crowding, it also *induces the investors exploiting* *the second characteristic to increase their investment positions, which reduces the negative impact of crowding in the first characteristic.*

5.3 Centralized investing in characteristics

We now consider a centralized setting in which a *single* investor exploits both characteristics. The following proposition shows that centralization leads to an increase in the total profits from exploiting both characteristics. The main takeaway from this result is that financial institutions have an incentive to *centralize* the exploitation of multiple characteristics because of trading diversification.

Proposition 5.5 *Let Assumptions 4.1 and 4.2 hold, then:*

1. *The equilibrium investment position in the first characteristic in the centralized setting, θ_{1d}, is larger than in the decentralized setting; that is, $\theta_{1c} > \theta_{1d} > 0$.*

2. *The equilibrium investment position in the second characteristic in the centralized setting, θ_{2c}, is negative and larger in absolute value than in the decentralized setting; that is, $\theta_{2c} < \theta_{2d} < 0$.*

3. *The profits from trading the second characteristic π_{2c} are zero in the centralized setting and strictly smaller than those in the decentralized setting; that is, $0 = \pi_{2c} < \pi_{2d}$.*

4. *The equilibrium total profits π_c and the equilibrium profits from the first characteristic in the centralized setting π_{1c} are larger than those in the decentralized setting; that is, $\pi_c > \pi_d$ and $\pi_{1c} > \pi_{1d}$.*

To understand the intuition underlying Proposition 5.5, note that centralizing the exploitation of two characteristics allows the single investor to internalize the three externalities present in the decentralized setting: among investors exploiting the first characteristic, among investors exploiting the second characteristic, and between the two groups of investors. After internalizing these externalities, the single investor makes decisions that maximize total profits. Another insight from Proposition 5.5 is that, for the case with pure liquidity-provision motive, the profits from the second characteristic are zero in the centralized setting.[15] That is, the second characteristic is used solely to increase the profit from exploiting the first characteristic.

6 Empirical calibration of game-theoretic model

To investigate the magnitude of the impact of trading diversification on the equilibrium, we now calibrate the game-theoretic model using historical return data and rebalancing-trade vectors along with the price-impact cost model of Frazzini et al. (2018), as in Section 3.2. We use "investment (asset growth)" as the first characteristic and "gross profitability" as the second characteristic. Section IA.1.2 in the Internet Appendix considers the case with a different characteristic, "book to market," as the first characteristic, but the same characteristic, "gross profitability," as the second. For the price-impact cost model of Frazzini et al. (2018), there are no closed-form expressions for the equilibrium quantities, so we compute these numerically.

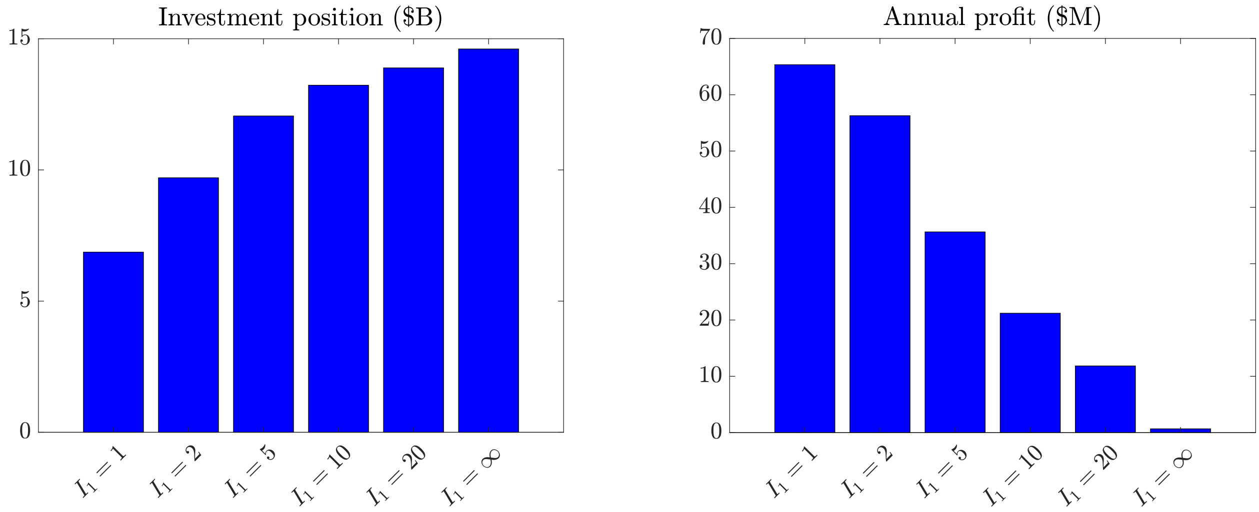
Figure 4 illustrates the effect of crowding on aggregate investment positions and profits when there are only investors exploiting the first characteristic ($I_1 \geq 1$ and $I_2 = 0$). The figure depicts the aggregate investment position and profits for the first characteristic for the cases with $I_1 = 1, 2, 5, 10, 20, \infty$ investors. Note that increasing the number of competing investors from one to twenty almost doubles the aggregate investment position and reduces the aggregate profits to less than a fifth because of crowding. In the limit as the investors become perfectly competitive ($I_1 = \infty$), their aggregate investment position is double that for the case with a single investor and their aggregate profits vanish.

Figure 5 depicts the investment positions and profits for the two characteristics for the decentralized setting where there is a single investor exploiting the first characteristic $I_1 = 1$

[15] To see this, note the first-order optimality conditions for the centralized setting portfolio problem (23) imply that $\lambda_2\theta_{2c} = -\lambda_{12}\theta_{1c}$; in other words, the price impact of the investment in the second characteristic cancels with the price impact of the interaction between the two characteristics.

Figure 4: Crowding with a single characteristic

This figure illustrates the effect of crowding on aggregate investment positions and profits when there are only investors exploiting the first characteristic ($I_1 \geq 1$ and $I_2 = 0$). The figure depicts the aggregate investment position and profits for the first characteristic for the cases with $I_1 = 1, 2, 5, 10, 20, \infty$ investors. We consider the price-impact cost model of Frazzini et al. (2018) and use "investment (asset growth)" as the first characteristic.

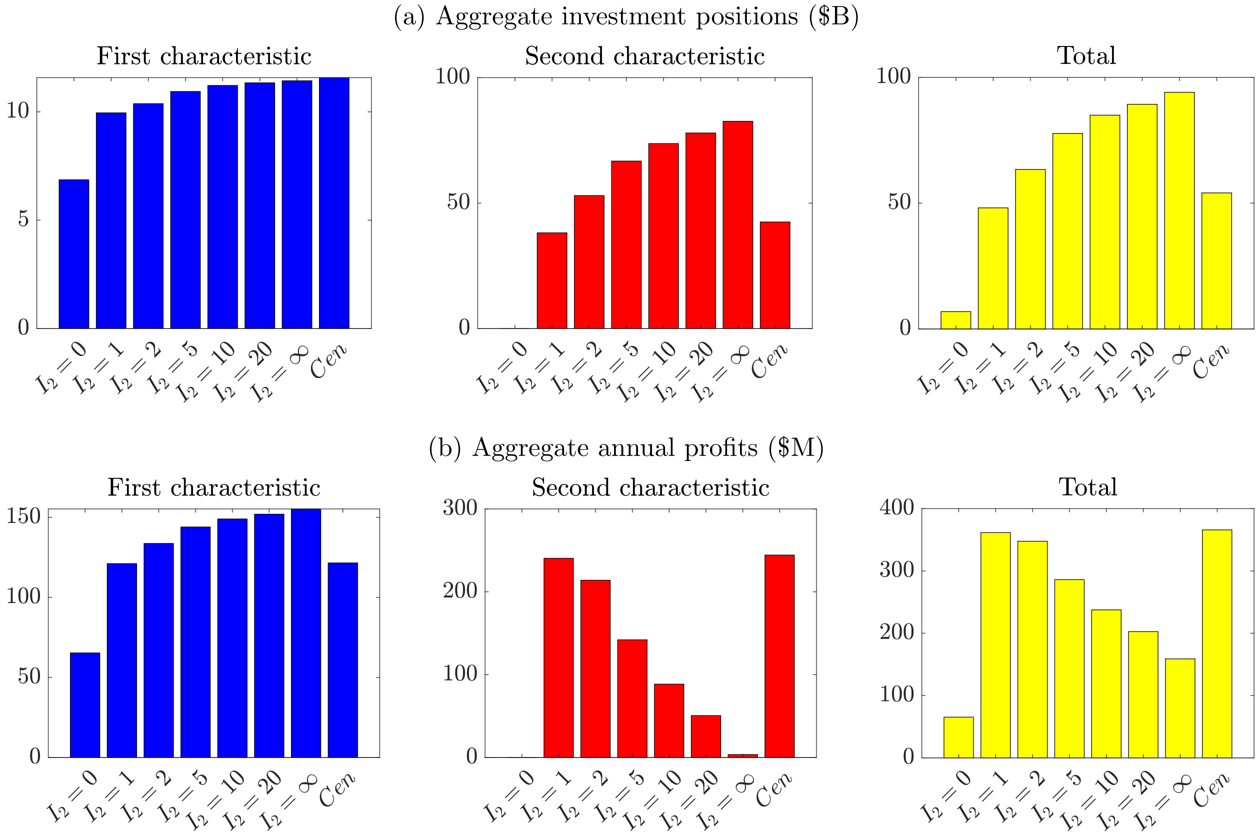


and $I_2 = 0, 1, 2, 5, 10, 20, \infty$ investors exploiting the second, and for the centralized setting (Cen). Panel (a) depicts the aggregate investment position in each of the two characteristics and the total investment position across both characteristics. Panel (b) depicts the profits obtained from each characteristic and the total profits from both characteristics.

Comparing the case where there are no investors ($I_2 = 0$) to the case where there is a single investor exploiting the second characteristic ($I_2 = 1$), we note from Figure 5 that trading diversification leads to a 45% increase in aggregate investment position and an 85% increase in profits from the first characteristic. Moreover, when the number of investors exploiting the second characteristic increases from one to twenty, the aggregate profits in the second characteristic are reduced to just above a fifth while the aggregate investment position in the second characteristic more than doubles. This additional investment in the second characteristic generates trading diversification benefits for the investors exploiting the first characteristic, who in response increase their aggregate investment by a further 14% and their aggregate profits by 25%. Overall, comparing the case without investors exploiting the second characteristic ($I_2 = 0$) to the case with twenty investors ($I_2 = 20$), trading diversification and competition among investors exploiting the second characteristic

Figure 5: Trading diversification and competition among investors in second characteristic

This figure depicts the investment positions and profits for the two characteristics for the decentralized setting where there is a single investor exploiting the first characteristic $I_1 = 1$ and $I_2 = 0, 1, 2, 5, 10, 20, \infty$ investors exploiting the second, and for the centralized setting (Cen). Panel (a) depicts the aggregate investment position in each of the two characteristics and the total investment position across both characteristics in billions of dollars. Panel (b) depicts the annual profits obtained from each characteristic and the total profits from both characteristics in millions of dollars. We consider the price-impact cost model of Frazzini et al. (2018) and use "investment (asset growth)" as the first characteristic and "gross profitability" as the second.



leads to a 65% increase in aggregate investment position and a 132% increase in aggregate profits from the first characteristic.

Finally, the single investor in the centralized setting maximizes the *total* profits across the two characteristics by taking an even greater investment position in the first characteristic, but a smaller position in the second characteristic, compared to the decentralized setting with $I_1 = 1$ and $I_2 = 20$. This is because by reducing the investment position in the second characteristic, the single investor substantially increases the profits from the second characteristic at the expense of only a modest reduction in the profits from the first characteristic, thus generating substantially higher total profits.

7 Conclusion

The explosion in the *number* of fund managers investing in factors has raised concerns about the effect of crowding on the profitability of these strategies. The analysis in our manuscript suggests that the answer to the question posed in the title is that the *trading-diversification* mechanism that we identify alleviates the effects of crowding in factor investing. That is, although competition among investors exploiting the *same* characteristic does erode their profits, competition among investors exploiting *different* characteristics *increases* the capacity and profits of factor-investing strategies due to trading diversification.

Our work has implications for various stakeholders in financial markets. First, financial institutions should search for characteristics that not only provide high returns, but are also exploited by a small number of competing institutions. Second, financial institutions should seek to exploit characteristics that allow them to benefit from the trading diversification generated by other institutions exploiting different characteristics. Third, regulators need to recognize that, although encouraging competition among fund managers exploiting a characteristic may reduce fees, it may also erode the profitability of factor-investing products because of crowding. However, encouraging the appropriate balance of competition between managers exploiting *different* characteristics can actually alleviate crowding and increase profits due to trading diversification.

A Proofs for all results

In this appendix we provide the proofs for all the results in the main body of the manuscript.

Proof of Proposition 3.1

For the case where the nth stock price-impact parameter is independently distributed from the rebalancing trades, the price-impact diversification ratio simplifies to

$$\text{price-impact diversification ratio} = \frac{E\left[\left|\sum_{k=1}^{K} \tilde{x}_{ktn}\right|^{1+\alpha}\right]}{\sum_{k=1}^{K} E\left[\left|\tilde{x}_{ktn}\right|^{1+\alpha}\right]}. \tag{A1}$$

Below, we characterize the expectation in the numerator and denominator on the right-hand side of Equation (A1) for α in the interval $[0, 1]$. Let

$$\tilde{x}_{tn}^{ew} = \sum_{k=1}^{K} \tilde{x}_{ktn}$$

be the trade in the nth stock required to rebalance an equally weighted portfolio of the K characteristics. Because $\tilde{x}_{ktn}$ for $k = 1, 2, \ldots, K$ are jointly distributed as a multivariate Normal distribution with zero mean and covariance matrix Ω, we have that $\tilde{x}_{tn}^{ew}$ is distributed as a Normal distribution with zero mean and variance $\sum_{k=1}^{K} \sigma_k^2 + \sum_{k=1}^{K} \sum_{l \neq k} \rho_{kl}\sigma_k\sigma_l$.

We need to characterize $E[|\tilde{x}_{tn}^{ew}|^{1+\alpha}]$. Because $\tilde{x}_{tn}^{ew}$ is distributed as a Normal distribution with zero mean, we have that $E[|\tilde{x}_{tn}^{ew}|^{1+\alpha}]$ is the central moment of order $1 + \alpha$ of a Normal random variable. Winkelbauer (2012) shows that for $\alpha > -1$

$$E[|\tilde{x}_{tn}^{ew}|^{1+\alpha}] = \frac{\Gamma(\frac{2+\alpha}{2})}{\pi} \times 2^{1+\alpha} \times \left(\sum_{k=1}^{K} \sigma_k^2 + \sum_{k=1}^{K} \sum_{l \neq k} \rho_{kl}\sigma_k\sigma_l\right)^{\frac{1+\alpha}{2}}, \tag{A2}$$

where $\Gamma(.)$ is the Gamma function; see Winkelbauer (2012, p. 1). Similarly, we have that

$$E[|\tilde{x}_{ktn}|^{1+\alpha}] = \frac{\Gamma(\frac{2+\alpha}{2})}{\pi} \times 2^{1+\alpha} \times \sigma_k^{1+\alpha}. \tag{A3}$$

Taking the ratio of (A2) to the summation of (A3) for $k = 1, 2, \ldots, K$, we get

$$\text{price-impact diversification ratio} = \frac{\left(\sum_{k=1}^{K} \sigma_k^2 + \sum_{k=1}^{K} \sum_{l \neq k} \rho_{kl}\sigma_k\sigma_l\right)^{\frac{1+\alpha}{2}}}{\sum_{k=1}^{K} \sigma_k^{1+\alpha}}. \tag{A4}$$

For the symmetric case with $\sigma_k^2 = \sigma^2$ for all k and $\rho_{kl} = \rho$ for all $k \neq l$, we have that

$$\text{price-impact diversification ratio} = \frac{\left[K(1+(K-1)\rho)\right]^{\frac{1+\alpha}{2}}}{K}, \tag{A5}$$

where the term $K(1+(K-1)\rho)$ is strictly positive because Ω is positive definite. Finally, the value of $\bar{\rho}$ follows using straightforward algebra.

Proof of Lemma 4.1

For the case with linear price impact, $\alpha = 1$, the price-impact at time t defined in Equation (3) becomes

$$\text{PI}_t = \Lambda_t \Delta w_t, \tag{A6}$$

where the aggregate amount of trading is

$$\Delta w_t = \sum_{k=1}^{2} \sum_{i=1}^{I_k} \Delta w_{kit}, \tag{A7}$$

in which Δw_{kit} contains the portfolio-rebalancing trades for the ith investor in the kth characteristic:

$$\Delta w_{kit} = w_{kit}(\theta_{ki}) - w_{kit}^{+}(\theta_{ki}), \tag{A8}$$

$w_{kit}^{+}(\theta_{ki})$ is the portfolio of the ith investor in the kth characteristic before trading at time t:

$$w_{kit}^{+}(\theta_{ki}) = \theta_{ki}\ x_{k,t-1} \circ (e + r_t), \tag{A9}$$

e is the N-dimensional vector of ones, and $x \circ y$ is the componentwise product of vectors x and y. The price-impact costs at time t of the ith investor in the kth characteristic is:

$$\text{PIC}_{kit} = \Delta w_{kit} \text{PI}_t.$$

The lemma follows from straightforward algebra.

Proof of Lemma 4.2

We first show the result for the empirically relevant case where there is a *discrete* joint probability distribution for the rebalancing-trade vectors, $\tilde{x}_{1t}$ and $\tilde{x}_{2t}$, and the price-impact matrix, Λ_t. By concatenating the rebalancing trades $\tilde{x}_{1t}$ and $\tilde{x}_{2t}$ and the matrices Λ_t for all realizations of the discrete distribution into panel vectors and matrices, it is straightforward to show that λ_1 and λ_2 are squared norms of certain vectors and λ_{12} is the scalar product of these same vectors. Therefore, it follows from the triangular inequality for norms that $\lambda_{12}^2 \leq \lambda_1\lambda_2$. Moreover, unless the rebalancing trades of the two characteristics are identical for every stock and every realization up to a change of scale, we have that the triangular inequality holds strictly $\lambda_{12}^2 < \lambda_1\lambda_2$. For the case where there is a continuous joint distribution for the rebalancing-trade vectors and the price-impact matrix, the result can be shown under mild assumptions by discretizing the continuous distribution and taking the limit when the granularity of the discretization goes to zero.

Proof of Proposition 4.1

Part 1. By Lemma 4.2 we know that $\lambda_k > 0$ for $k = 1, 2$ and thus, the decision problem of the ith investor in the kth characteristic is *strictly* convex. Therefore, there exists a unique global minimizer to the decision problem of the ith investor in the kth characteristic and it is given by the solution to the first-order optimality conditions:

$$2\lambda_1\theta_{1i} + \lambda_1\theta_{1,-i} + \lambda_{12}\sum_{j=1}^{I_2}\theta_{2j} = \mu_1, \quad \text{and} \tag{A10}$$

$$2\lambda_2\theta_{2i} + \lambda_2\theta_{2,-i} + \lambda_{12}\sum_{j=1}^{I_1}\theta_{1j} = \mu_2. \tag{A11}$$

Therefore, the investment positions θ_{1id} and θ_{2id} are a Nash equilibrium if and only if they satisfy the first-order optimality conditions of the investors in the first and second characteristics; that is, if they satisfy the following system of linear equations:

$$\begin{pmatrix} 2\lambda_1 & \lambda_1 & \cdots & \lambda_1 & \lambda_{12} & \lambda_{12} & \cdots & \lambda_{12} \\ \lambda_1 & 2\lambda_1 & \cdots & \lambda_1 & \lambda_{12} & \lambda_{12} & \cdots & \lambda_{12} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_1 & \lambda_1 & \cdots & 2\lambda_1 & \lambda_{12} & \lambda_{12} & \cdots & \lambda_{12} \\ \lambda_{12} & \lambda_{12} & \cdots & \lambda_{12} & 2\lambda_2 & \lambda_2 & \cdots & \lambda_2 \\ \lambda_{12} & \lambda_{12} & \cdots & \lambda_{12} & \lambda_2 & 2\lambda_2 & \cdots & \lambda_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_{12} & \lambda_{12} & \cdots & \lambda_{12} & \lambda_2 & \lambda_2 & \cdots & 2\lambda_2 \end{pmatrix} \begin{pmatrix} \theta_{11d} \\ \theta_{12d} \\ \vdots \\ \theta_{1I_1d} \\ \theta_{21d} \\ \theta_{22d} \\ \vdots \\ \theta_{2I_2d} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_1 \\ \vdots \\ \mu_1 \\ \mu_2 \\ \mu_2 \\ \vdots \\ \mu_2 \end{pmatrix}. \tag{A12}$$

We now prove that there is a unique Nash equilibrium by showing that the matrix on the left hand side of (A12) is nonsingular. Assume by contradiction that there is a nonzero vector of $\theta_{1i}'s$ and θ_{2i}'s that satisfies the following:

$$\begin{pmatrix} 2\lambda_1 & \lambda_1 & \cdots & \lambda_1 & \lambda_{12} & \lambda_{12} & \cdots & \lambda_{12} \\ \lambda_1 & 2\lambda_1 & \cdots & \lambda_1 & \lambda_{12} & \lambda_{12} & \cdots & \lambda_{12} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_1 & \lambda_1 & \cdots & 2\lambda_1 & \lambda_{12} & \lambda_{12} & \cdots & \lambda_{12} \\ \lambda_{12} & \lambda_{12} & \cdots & \lambda_{12} & 2\lambda_2 & \lambda_2 & \cdots & \lambda_2 \\ \lambda_{12} & \lambda_{12} & \cdots & \lambda_{12} & \lambda_2 & 2\lambda_2 & \cdots & \lambda_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_{12} & \lambda_{12} & \cdots & \lambda_{12} & \lambda_2 & \lambda_2 & \cdots & 2\lambda_2 \end{pmatrix} \begin{pmatrix} \theta_{11} \\ \theta_{12} \\ \vdots \\ \theta_{1I_1} \\ \theta_{21} \\ \theta_{22} \\ \vdots \\ \theta_{2I_2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{A13}$$

Then, any solution to (A13) must satisfy the first I_1 equations in (A13), which can be rewritten as

$$\begin{pmatrix} 2\lambda_1 & \lambda_1 & \cdots & \lambda_1 \\ \lambda_1 & 2\lambda_1 & \cdots & \lambda_1 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 & \lambda_1 & \cdots & 2\lambda_1 \end{pmatrix} \begin{pmatrix} \theta_{11} \\ \theta_{12} \\ \vdots \\ \theta_{1I_1} \end{pmatrix} = -\lambda_{12} \sum_{i=1}^{I_2} \theta_{2i} e, \tag{A14}$$

where e is the I_1-dimensional vector of ones. The matrix on the left-hand side of (A14) is nonsingular because by Lemma 4.2 we know that $\lambda_1 > 0$. Moreover, this matrix is symmetric with respect to the I_1 investors in the first characteristic. Therefore, any solution to Equation (A14) must be symmetric with respect to the I_1 investors in the first characteristic;

that is, $\theta_{1i} = \theta_1$ for $i = 1,\, 2,\, \ldots, I_1$. Consequently Equation (A13) can be rewritten as

$$\left(\begin{array}{ccccc} (I_1+1)\lambda_1 & \lambda_{12} & \lambda_{12} & \cdots & \lambda_{12} \\ I_1\lambda_{12} & 2\lambda_2 & \lambda_2 & \cdots & \lambda_2 \\ I_1\lambda_{12} & \lambda_2 & 2\lambda_2 & \cdots & \lambda_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ I_1\lambda_{12} & \lambda_2 & \lambda_2 & \cdots & 2\lambda_2 \end{array}\right) \left(\begin{array}{c} \theta_1 \\ \theta_{21} \\ \theta_{22} \\ \vdots \\ \theta_{2I_2} \end{array}\right) = \left(\begin{array}{c} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{array}\right). \tag{A15}$$

Using similar arguments as above, it is easy to show that any solution to Equation (A15) must be symmetric with respect to the I_2 investors in the second characteristic; that is, $\theta_{2d} = \theta_2$ for $i = 1,\, 2, \ldots,\, I_2$. Thus, we can express (A13) as follows

$$\left(\begin{array}{cc} (I_1+1)\lambda_1 & I_2\lambda_{12} \\ I_1\lambda_{12} & (I_2+1)\lambda_2 \end{array}\right) \left(\begin{array}{c} \theta_1 \\ \theta_2 \end{array}\right) = \left(\begin{array}{c} 0 \\ 0 \end{array}\right). \tag{A16}$$

The matrix on the left-hand side of (A16) is nonsingular for any I_1 and I_2 different from zero because its determinant is $(I_1+1)(I_2+1)\lambda_1\lambda_2 - I_1 I_2 \lambda_{12}^2$, which is nonzero by Lemma 4.2. Consequently, there is a unique Nash equilibrium given by the unique solution to the linear system of equations in (A12).

Part 2. By arguments similar to those in Part 1, any solution to (A12) must be symmetric with respect to the I_1 investors in the first characteristic and with respect to the I_2 investors in the second characteristic; that is, $\theta_{kid} = \theta_{kd}$ for $i = 1, 2, \ldots, I_k$ and $k = 1, 2$.

Part 3. Therefore, the unique equilibrium is the solution to the following system of two linear equations with two variables

$$\left(\begin{array}{cc} (I_1+1)\lambda_1 & I_2\lambda_{12} \\ I_1\lambda_{12} & (I_2+1)\lambda_2 \end{array}\right) \left(\begin{array}{c} \theta_{1d} \\ \theta_{2d} \end{array}\right) = \left(\begin{array}{c} \mu_1 \\ \mu_2 \end{array}\right). \tag{A17}$$

The above system of two equations can be solved by premultiplying the vector of characteristic means by the inverse of the left-hand side matrix. This gives the following optimal solutions:

$$\theta_{1id} = \frac{(I_2+1)\,\lambda_2\mu_1 - I_2\lambda_{12}\mu_2}{(I_1+1)(I_2+1)\lambda_1\lambda_2 - I_1 I_2 \lambda_{12}^2},$$

$$\theta_{2id} = \frac{(I_1+1)\lambda_1\mu_2 - I_1\lambda_{12}\mu_1}{(I_1+1)(I_2+1)\lambda_1\lambda_2 - I_1 I_2 \lambda_{12}^2}.$$

Part 4. The profit of the ith investor in the kth characteristic is her expected return net of price impact multiplied by her investment position. Therefore, it suffices to show that the expected return net of price impact of the ith investor in the kth characteristic is $\bar{\mu}_{kid} = \lambda_k \theta_{kid}$. The expected return net of price impact of the ith investor in the first characteristic is

$$\bar{\mu}_{1id} = \mu_1 - \lambda_1 I_1 \theta_{1id} - \lambda_{12} I_2 \theta_{2id}.$$

Now, using the ith investor's first-order conditions, we have that:

$$0 = \mu_1 - \lambda_1 (I_1 + 1)\theta_{1id} - \lambda_{12} I_2 \theta_{2id}.$$

Therefore, substituting the last equation into the expression for $\bar{\mu}_{1id}$, we obtain $\bar{\mu}_{1id} = \lambda_1 \theta_{1id}$. The result for the ith investor in the second characteristic is obtained similarly.

Proof of Proposition 4.2

Part 1. The decision in the centralized setting is given in (23). By Lemma 4.2 we have that $\lambda_1 \lambda_2 > \lambda_{12}^2$ and therefore the decision problem in the centralized setting is strictly convex and there exists a unique minimizer.

Part 2. The unique minimizer is given by the first-order optimality conditions for the single investor in the centralized setting:

$$\begin{pmatrix} 2\lambda_1 & 2\lambda_{12} \\ 2\lambda_{12} & 2\lambda_2 \end{pmatrix} \begin{pmatrix} \theta_{1c} \\ \theta_{2c} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}. \tag{A18}$$

The result follows from straightforward algebra.

Proof of Proposition 5.1

Note that the decision problems of the ith investor in the first characteristic in the *absence* and *presence* of investors in the second characteristic are identical for the case with $\lambda_{12} = 0$. Therefore, the equilibrium investment position and profits of the ith investor in the first characteristic in the absence of investors in the second characteristic are obtained by setting $\lambda_{12} = 0$ in Equations (24) and (26) of Proposition 4.1.

The monotonicity results follow from Equations (31) and (32) by noting that $I_1/(I_1+1)$ is increasing and $I_1/(I_1+1)^2$ is decreasing in I_1 for all $I_1 \geq 1$

Proof of Proposition 5.2

To obtain the capacity of the first characteristic, we first determine the best response of the investors in the second characteristic to a given aggregate investment position in the first characteristic θ_{1d}. Note that the decision problem of the ith investor in the second characteristic for the case with pure liquidity-provision motive is

$$\min_{\theta_{2i}} \quad \theta_{2i}\lambda_2(\theta_{2i} + \theta_{2,-i}) + \theta_{2i}\lambda_{12}\theta_{1d}.$$

Thus, the first-order optimality condition for the ith investor in the second characteristic is

$$2\lambda_2\theta_{2i} + \lambda_2\theta_{2,-i} = -\lambda_{12}\theta_{1d}.$$

It follows from the proof of Proposition 4.1 that the equilibrium among investors in the second characteristic is symmetric, and thus we can rewrite the first-order optimality conditions as

$$(I_2 + 1)\lambda_2\theta_{2i} = -\lambda_{12}\theta_{1d},$$

and therefore the aggregate best response of the investors in the second characteristic is

$$\theta_{2d} = -\frac{I_2}{I_2+1}\frac{\lambda_{12}}{\lambda_2}\theta_{1d}. \tag{A19}$$

The capacity of the first characteristic is the aggregate investment position in the first characteristic for which its aggregate profits are zero, which must satisfy the following equation:

$$\theta_{1d}\lambda_1\theta_{1d} + \theta_{1d}\lambda_{12}\theta_{2d} - \theta_{1d}\mu_1 = 0.$$

We can simplify this equation by removing the trivial root $\theta_{1d} = 0$ and we obtain

$$\lambda_1\theta_{1d} + \lambda_{12}\theta_{2d} - \mu_1 = 0.$$

Plugging (A19) into this equation we obtain that the capacity of the first characteristic is

$$\theta_{1d} = \frac{\mu_1}{\lambda_1 - \frac{I_2}{I_2+1}\frac{\lambda_{12}^2}{\lambda_2}}.$$

Proof of Proposition 5.3

Part 1. The partial derivative of the aggregate investment position in the first characteristic with respect to I_2 is

$$\frac{\partial(\theta_{1d})}{\partial I_2} = \frac{\partial(I_1\theta_{1id})}{\partial I_2} = I_1\frac{\partial(\theta_{1id})}{\partial I_1}$$

$$= I_1\frac{\lambda_2\mu_1\Big((I_1+1)(I_2+1)\lambda_1\lambda_2 - I_1I_2\lambda_{12}^2\Big) - (I_2+1)\lambda_2\mu_1\Big((I_1+1)\lambda_1\lambda_2 - I_1\lambda_{12}^2\Big)^2}{\Big((I_1+1)(I_2+1)\lambda_1\lambda_2 - I_1I_2\lambda_{12}^2\Big)^2}$$

$$= I_1\lambda_2\mu_1\frac{\Big((I_1+1)(I_2+1)\lambda_1\lambda_2 - I_1I_2\lambda_{12}^2\Big) - (I_2+1)\Big((I_1+1)\lambda_1\lambda_2 - I_1\lambda_{12}^2\Big)^2}{\Big((I_1+1)(I_2+1)\lambda_1\lambda_2 - I_1I_2\lambda_{12}^2\Big)^2} > 0,$$

where the last inequality follows from the fact that the ratio is strictly positive for all $I_2 > 0$ because of Lemma 4.2.

Part 2. The partial derivative of the aggregate investment position in the second characteristic with respect to I_2 is

$$\frac{\partial(\theta_{2d})}{\partial I_2} = \frac{\partial(I_2\theta_{2id})}{\partial I_2} = \theta_{2id} + I_2\frac{\partial(\theta_{2id})}{\partial I_2} \tag{A20}$$

$$= \theta_{2id}\left(1 - \frac{(I_1+1)\lambda_1\lambda_2 - I_1\lambda_{12}^2}{(I_1+1)\frac{I_2+1}{I_2}\lambda_1\lambda_2 - I_1\lambda_{12}^2}\right) < 0, \tag{A21}$$

where the inequality in (A21) holds because the ratio inside the parenthesis is strictly greater than one because of Lemma 4.2 and the fact that $(I_2+1)/I_2 > 1$ for finite I_2.

Part 3. The result follows from Part 1 and Equation (26).

Part 4. The partial derivative of the aggregate profit from the second characteristic with respect to I_2 is

$$\frac{\partial I_2\pi_{2id}}{\partial I_2} = \lambda_2\theta_{2id}^2\left(1 - 2\frac{(I_2+1)\lambda_1\lambda_2 - I_1\lambda_{12}^2}{(I_1+1)\frac{I_2+1}{I_2}\lambda_1\lambda_2 - I_1I_2\lambda_{12}^2}\right) < 0, \tag{A22}$$

where the inequality in (A22) follows from the fact that by Lemma 4.2, the ratio inside the parenthesis is greater than one provided $\left(2 - \frac{I_2+1}{I_2}\right) > \frac{I_1}{I_1+1}$.

Proof of Proposition 5.4

Part 1. The partial derivative of the aggregate investment position in the first characteristic with respect to I_1 is

$$\frac{\partial(\theta_{1d})}{\partial I_1} = \frac{\partial(I_1\theta_{1id})}{\partial I_1} = \theta_{1id} + I_1\frac{\partial(\theta_{1id})}{\partial I_1}$$

$$= \theta_{1id}\left(1 - I_1\frac{(I_2+1)\lambda_1\lambda_2 - I_2\lambda_{12}^2}{(I_2+1)(I_1+1)\lambda_1\lambda_2 - I_1I_2\lambda_{12}^2}\right)$$

$$= \theta_{1id}\left(1 - \frac{(I_2+1)\lambda_1\lambda_2 - I_2\lambda_{12}^2}{(I_2+1)\frac{I_1+1}{I_1}\lambda_1\lambda_2 - I_2\lambda_{12}^2}\right) > 0, \tag{A23}$$

where the inequality in (A23) follows because the ratio in the second term inside the parenthesis is positive and strictly smaller than one because of Lemma 4.2 and the fact that $(I_1+1)/I_1 > 1$ for finite I_1.

Part 2. When $\lambda_{12} > 0$, dividing by I_1 the numerator and denominator of the optimal investment position in the second characteristic, it is straightforward to see that the denominator becomes smaller as I_1 increases, whereas the numerator is independent of I_1 and always negative under Assumption 4.2 and given $\lambda_{12} > 0$. The overall result of these two effects is that the optimal investment position decreases with I_1 when $\lambda_{12} > 0$.

Part 3. The partial derivative of the aggregate profit from the first characteristic with respect to I_1 is

$$\frac{\partial(I_1\pi_{1id})}{\partial I_1} = \frac{\partial(I_1\lambda_1\theta_{1id}^2)}{\partial I_1} = \lambda_1\frac{\partial(I_1\theta_{1id}^2)}{\partial I_1}$$

$$= \lambda_1\left(\theta_{1id}^2 + 2I_1\theta_{1id}\frac{\partial(\theta_{1id})}{\partial I_1}\right). \tag{A24}$$

Plugging the partial derivative of θ_{1id} with respect to I_1 into (A24), we then have that

$$\frac{\partial(I_1\pi_{1id})}{\partial I_1} = \lambda_1\theta_{1id}^2\left(1 - \frac{2I_1((I_2+1)\lambda_1\lambda_2 - I_2\lambda_{12}^2)}{(I_2+1)(I_1+1)\lambda_1\lambda_2 - I_1I_2\lambda_{12}^2}\right)$$

$$= \lambda_1\theta_{1id}^2\left(1 - \frac{2I_1((I_2+1)\lambda_1\lambda_2 - I_2\lambda_{12}^2)}{(I_1+1)\left((I_2+1)\lambda_1\lambda_2 - \frac{I_1I_2}{I_1+1}\lambda_{12}^2\right)}\right). \tag{A25}$$

The ratio inside the parenthesis in (A25) is greater than one iff:

$$2I_1((I_2+1)\lambda_1\lambda_2 - I_2\lambda_{12}^2) > (I_1+1)\bigg((I_2+1)\lambda_1\lambda_2 - \frac{I_1 I_2}{I_1+1}\lambda_{12}^2\bigg). \tag{A26}$$

Simplifying this inequality we get

$$\bigg(2I_1(I_2+1) - (I_1+1)(I_2+1)\bigg)\lambda_1\lambda_2 > I_1 I_2 \lambda_{12}^2, \tag{A27}$$

which holds for all I_1 such that $\frac{I_1-1}{I_1} > \frac{I_2}{I_2+1}$. Thus, $\frac{\partial(I_1\pi_{1id})}{\partial I_1} > 0$ for all I_1 such that $\frac{I_1-1}{I_1} > \frac{I_2}{I_2+1}$.

Part 4. This result can be proven by using similar arguments to those in Part 2.

Proof of Proposition 5.5

Part 1. To show that the investment position in the first characteristic in the decentralized setting with $I_1 = 1$ and pure liquidity provision ($\mu_2 = 0$) is smaller than that of the centralized setting, we need to prove the following inequality:

$$\underbrace{\frac{2\lambda_2\mu_1}{4\lambda_1\lambda_2 - \lambda_{12}^2}}_{\theta_{1d}} < \underbrace{\frac{\lambda_2\mu_1}{2(\lambda_1\lambda_2 - \lambda_{12}^2)}}_{\theta_{1c}}. \tag{A28}$$

Simplifying we have

$$\frac{1}{4\lambda_1\lambda_2 - \lambda_{12}^2} < \frac{1}{4\lambda_1\lambda_2 - 4\lambda_{12}^2}.$$

By Lemma 4.2, we know that the denominators of both the right- and left-hand sides of the inequality are strictly positive. Also, the denominator of the right-hand side term is smaller and thus Inequality (A28) holds.

Part 2. We now prove that the investment position in the second characteristic in the decentralized setting with $I_1 = 1$ and pure liquidity provision ($\mu_2 = 0$) is negative but higher than that in the centralized setting. Therefore, we prove the following inequalities:

$$0 > \underbrace{\frac{-\lambda_{12}\mu_1}{4\lambda_1\lambda_2 - \lambda_{12}^2}}_{\theta_{2d}} > \underbrace{\frac{-\lambda_{12}\mu_1}{2(\lambda_1\lambda_2 - \lambda_{12}^2)}}_{\theta_{2c}}. \tag{A29}$$

Under Assumption 4.2, for the nontrivial case with $\lambda_{12} > 0$, we have that the numerators of θ_{2d} and θ_{2c} are identical and negative, whereas the denominators of θ_{2d} and θ_{2c} are strictly positive by Lemma 4.2. However, the denominator of θ_{2d} is larger than that of θ_{2c}, and thus, θ_{2d} is smaller in absolute value than θ_{2c}.

Parts 3. From Equation (30), we know that the profits from the second characteristic in the centralized setting are zero for the case with pure liquidity-provision motive, $\mu_2 = 0$. Moreover, Part 2 above and Equation (26) imply that the profits from the second characteristic in the decentralized setting are strictly positive.

Parts 4. The total profits in the decentralized setting have to be smaller than those in the centralized setting because by Proposition 4.2 we know that the optimal investment positions in the centralized setting are the unique minimizer to the total profit function.

Because we know from Part 3 that profits from the second characteristic are smaller in the centralized setting, and we have just shown that total profits are higher in the centralized setting, then we must have that profits from the first characteristic are larger in the centralized setting.

References

Almgren, Robert, Chee Thum, Emmanuel Hauptmann, and Hong Li, 2005, Direct estimation of equity market impact, *Risk* 18, 58–62.

Baert, Rick, 2018, Growth of ETFs reflects passive shift; 3 largest firms hold 79% of assets, *Pensions & Investments* (May 28).

Barroso, Pedro, and Pedro Santa-Clara, 2015, Beyond the carry trade: Optimal currency portfolios, *Journal of Financial and Quantitative Analysis* 50, 1037–1056.

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

Berk, Jonathan B., and Jules H. van Binsbergen, 2017, Mutual funds in equilibrium, *Annual Review of Financial Economics* 9, 147–167.

Bonelli, Maxime, Augustin Landier, Guillaume Simon, and David Thesmar, 2019, The capacity of trading strategies, available at SSRN 2585399.

Brandt, Michael W., Pedro Santa-Clara, and Rossen Valkanov, 2009, Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns, *Review of Financial Studies* 22, 3411–3447.

Carlson, Debbie, 2019, Invesco focusing on scale, *ETF.com* (January 24).

Cournot, Antoine-Augustin, 1838, *Recherches sur les Principes Mathématiques de la Théorie des Richesses par Augustin Cournot* (chez L. Hachette, Paris).

DeMiguel, Victor, Alberto Martin-Utrera, Francisco J. Nogales, and Raman Uppal, 2020, A transaction-cost perspective on the multitude of firm characteristics, *Review of Financial Studies* 33, 2180–2222.

Edelen, Roger M., Richard B. Evans, and Gregory B. Kadlec, 2007, Scale effects in mutual fund performance: The role of trading costs, available at SSRN 951367.

Feldman, David, Konark Saxena, and Jingrui Xu, 2018, Is the active fund management industry concentrated enough?, forthcoming in *Journal of Financial Economics*.

Feldman, David, Konark Saxena, and Jingrui Xu, 2019, The international active fund management industry: Concentration cross effects, available at SSRN 3330131.

Flood, Chris, 2019, Investors' smart money piles into smart beta ETFs, *Financial Times* (February 11).

Frazzini, Andrea, Ronen Israel, and Tobias J. Moskowitz, 2015, Trading costs of asset pricing anomalies, Fama-Miller Working Paper.

Frazzini, Andrea, Ronen Israel, and Tobias J. Moskowitz, 2018, Trading costs, AQR Working Paper.

Green, Jeremiah, John R.M. Hand, and X. Frank Zhang, 2017, The characteristics that provide independent information about average U.S. monthly stock returns, *Review of Financial Studies* 30, 4389–4436.

Grinold, Richard C., and Ronald N. Kahn, 2000, *Active portfolio management* (McGraw-Hill, New York), 2nd edition.

Grossman, Sanford J., and Joseph E. Stiglitz, 1980, On the impossibility of informationally efficient markets, *American Economic Review* 70, 393–408.

Harvey, Campbell R., Yan Liu, Eric K.M. Tan, and Min Zhu, 2020, Crowding: Evidence from fund managerial structure, available at SSRN 3554636.

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

Hoberg, Gerard, Nitin Kumar, and Nagpurnanand Prabhala, 2019, Buy-side competition and momentum profits, available from SSRN 3132378.

Jacobs, Bruce I., and Kenneth N. Levy, 2014, Investing in a multidimensional market, *Financial Analysts Journal* 70, 6–12.

Korajczyk, Robert A., and Ronnie Sadka, 2004, Are momentum profits robust to trading costs?, *Journal of Finance* 59, 1039–1082.

Kyle, Albert S, 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315–1335.

Novy-Marx, Robert, and Mihail Velikov, 2016, A taxonomy of anomalies and their trading costs, *Review of Financial Studies* 29, 104–147.

Pástor, Ľuboš, and Robert F. Stambaugh, 2012, On the size of the active management industry, *Journal of Political Economy* 120, 740–781.

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

Ratcliffe, Ronald, Paolo Miranda, and Andrew Ang, 2017, Capacity of smart beta strategies: A transaction cost perspective, *Journal of Index Investing* 8, 39–50.

Riding, Siobhan, 2018, Smart beta moves into mainstream for large investors, *Financial Times* (November 5).

Stein, Jeremy C., 2009, Presidential address: Sophisticated investors and market efficiency, *Journal of Finance* 64, 1517–1548.

Torre, Nicolo G., and Mark J. Ferrari, 1997, *Market Impact Model Handbook* (BARRA Inc, Berkeley).

Wahal, Sunil, and Albert Y. Wang, 2011, Competition among mutual funds, *Journal of Financial Economics* 99, 40–59.

Winkelbauer, Andreas, 2012, Moments and absolute moments of the normal distribution, arXiv preprint arXiv:1209.4340.

Internet Appendix to

What Alleviates Crowding in Factor Investing?

IA.1 Extensions to the base-case model

In this appendix we first show that the findings from our game-theoretic model in Section 4 are robust to considering investors that are risk averse rather than risk neutral. We then report the results from the empirical calibration in Section 6 for the case in which we use book to market (instead of asset growth) as the first characteristic.

IA.1.1 Risk-averse investors

In the main body of the manuscript, we consider risk-neutral investors. We now extend the model to study the robustness of our results to considering risk-averse investors. We assume that the absolute risk-aversion parameters of the investors in the first and second characteristics in the decentralized setting are $\gamma_1 = \frac{I_1+1}{2}\overline{\gamma}_1$ and $\gamma_2 = \frac{I_2+1}{2}\overline{\gamma}_2$, respectively. This is reasonable because each investor makes a smaller investment as the number of competing investors increases, and hence their absolute risk aversion must increase with the number of competing investors.

The ith investor in the first characteristic chooses her investment position θ_{1i} to optimize her mean-variance utility net of price-impact costs

$$\min_{\theta_{1i}} \quad \frac{\gamma_1}{2}\theta_{1i}\sigma_1^2\theta_{1i} + \theta_{1i}\lambda_1(\theta_{1i} + \theta_{1,-i}) + \theta_{1i}\lambda_{12}\sum_{j=1}^{I_2}\theta_{2j} - \theta_{1i}\mu_1, \tag{IA.1.1}$$

where σ_1^2 is the variance of the first characteristic return. Similarly, the decision problem of the ith investor in the second characteristic is

$$\min_{\theta_{2i}} \quad \frac{\gamma_2}{2}\theta_{2i}\sigma_2^2\theta_{2i} + \theta_{2i}\lambda_2(\theta_{2i} + \theta_{2,-i}) + \theta_{2i}\lambda_{12}\sum_{j=1}^{I_1}\theta_{1j} - \theta_{2i}\mu_2, \tag{IA.1.2}$$

where σ_2^2 is the variance of the second characteristic return. Using similar arguments to those in the proofs of Propositions 4.1 and 4.2, the equilibrium is symmetric and thus the optimality condition of the ith investor in the first characteristic can be written as

$$\gamma_1\sigma_1^2\theta_{1i} + (I_1+1)\lambda_1\theta_{1i} + I_2\lambda_{12}\theta_{2j} - \mu_1 = 0, \tag{IA.1.3}$$

which can be rewriten as

$$(I_1+1)\left(\frac{\overline{\gamma}_1}{2}\sigma_1^2 + \lambda_1\right)\theta_{1i} + I_2\lambda_{12}\theta_{2j} - \mu_1 = 0. \tag{IA.1.4}$$

Similarly, the optimality condition for the ith investor in the second characteristic is

$$(I_2 + 1)\left(\frac{\overline{\gamma}_2}{2}\sigma_2^2 + \lambda_2\right)\theta_{2i} + I_1\lambda_{12}\theta_{1j} - \mu_2 = 0. \tag{IA.1.5}$$

From these optimality conditions, it is straightforward to show that the equilibrium quantities in the case with risk-averse investors are those given in Propositions 4.1 and 4.2 of the main body of the manuscript after replacing the transaction cost parameters λ_1 and λ_2 with $\widetilde{\lambda}_1 = \frac{\overline{\gamma}_1}{2}\sigma_1^2 + \lambda_1$ and $\widetilde{\lambda}_2 = \frac{\overline{\gamma}_2}{2}\sigma_2^2 + \lambda_2$, respectively. Therefore, the results in the main body of the manuscript continue to hold for the case with risk-averse investors.

IA.1.2 Empirical calibration: book to market and profitability

In Section 6, we calibrate the game-theoretic model for the case with "investment (asset growth)" as the first characteristic and "gross profitability" as the second. We now calibrate the game-theoretic model with a different first characteristic, "book to market," as the first characteristic, and the same characteristic, "gross profitability," as the second. Figure IA.1 depicts the investment positions and profits when there are $I_1 = 1, 2, 5, 10,\ 20, \infty$ investors exploiting the first characteristic in the absence of investors exploiting the second ($I_2 = 0$). Figure IA.2 depicts the investment positions and profits for the decentralized setting with $I_1 = 1$ investor in the first characteristic and $I_2 = 0, 1, 2, 5, 10, 20, \infty$ investors in the second, and for the centralized setting (Cen). The results for the case with "book to market" as the first characteristic are similar to those obtained in Section 6 for the case with "investment (asset growth)" as the first characteristic. In particular, Figure IA.1 shows that competition among investors exploiting the first characteristic erodes their aggregate profits due to crowding. Figure IA.2 shows that trading diversification and competition among investors exploiting the second characteristic alleviate crowding in the first characteristic.

Figure IA.1: Crowding with "book to market" as the single characteristic

This figure illustrates the effect of crowding on aggregate investment positions and profits when there are only investors exploiting the first characteristic ($I_1 \geq 1$ and $I_2 = 0$). The figure depicts the aggregate investment position and profits for the first characteristic for the cases with $I_1 = 1, 2, 5, 10, 20, \infty$ investors. We consider the price-impact cost model of Frazzini et al. (2018) and use "book to market" as the first characteristic.

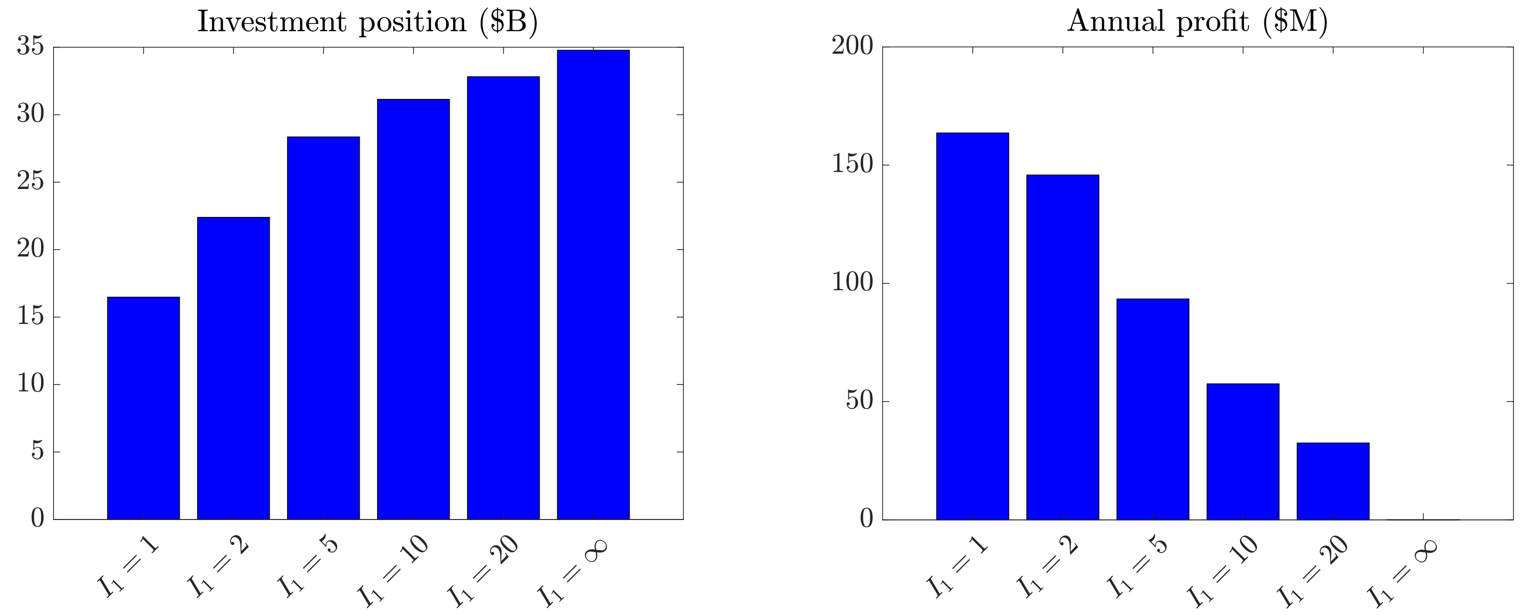


Figure IA.2: Trading diversification and competition with "book to market" as first characteristic

This figure depicts the investment positions and profits for the two characteristics for the decentralized setting where there is a single investor exploiting the first characteristic $I_1 = 1$ and $I_2 = 0, 1, 2, 5, 10, 20, \infty$ investors exploiting the second, and for the centralized setting (Cen). Panel (a) depicts the aggregate investment position in each of the two characteristics and the total investment position across both characteristics in billions of dollars. Panel (b) depicts the annual profits obtained from each characteristic and the total profits from both characteristics in millions of dollars. We consider the price-impact cost model of Frazzini et al. (2018) and use "book to market" as the first characteristic and "gross profitability" as the second.