Online Appendix for Quantifying the Impact of Impact Investing

Andrew W. Lo* and Ruixun Zhang†

9 Aug 2023

Contents

*Charles E. and Susan T. Harris Professor, MIT Sloan School of Management; Director, MIT Laboratory for Financial Engineering; Principal Investigator, MIT Computer Science and Artificial Intelligence Laboratory; External Faculty, Santa Fe Institute.

†Assistant Professor and Boya Young Fellow, Peking University School of Mathematical Sciences, Center for Statistical Science, National Engineering Laboratory for Big Data Analysis and Applications, and Laboratory for Mathematical Economics and Quantitative Finance. zhangruixun@pku.edu.cn (email).

A Additional Technical Results

In this Appendix, we provide additional technical results.

A.1 Asymptotic Distribution of Induced Order Statistics

As the number of securities, N, increases without bound, the limiting joint distribution of the induced order statistics, $\alpha_{[i:N]}$, has been derived by Yang (1977) and does not require the normality assumption (A1), hence we can rely on this asymptotic approximation for large samples.

Proposition A.1. *Assuming $(X_1, \alpha_1)^T, \cdots, (X_N, \alpha_N)^T$ are IID, for any sequence $1 < i_1 < \cdots < i_n < N$ such that, as $N \to \infty$, $i_k/N \to \xi_k \in (0,1)$ for $k = 1, \cdots, n$, we have:*

$$\lim_{N\to\infty} \mathbb{P}\left(\alpha_{[i_1:N]} < a_1, \cdots, \alpha_{[i_n:N]} < a_n\right) = \prod_{k=1}^{n} \mathbb{P}\left(\alpha_k < a_k | F_x(X_k) = \xi_k\right), \tag{A.1}$$

where $F_x(\cdot)$ is the marginal CDF of X_i.

Proposition A.1 implies that the induced order statistics at distinct quantiles are asymptotically independent, consistent with the finite sample observations in Proposition 3 and Figure 1. Also, because the conditional distribution of jointly normal random vectors is still normal, we can characterize the first two moments of the induced order statistics asymptotically via the following result.

Proposition A.2. *Under Assumption (A1), as N increases without bound, the induced order statistics, $\alpha_{[i_k:N]}$ $(k = 1, \cdots, n)$, converge in distribution to independent Gaussian random variables with mean $\mu(\xi_k)$ and variance $\sigma^2(\xi_k)$, where*

$$\mu(\xi_k) \equiv \rho(\sigma_\alpha/\sigma_x)\left[F_x^{-1}(\xi_k) - \mu_x\right] = \rho\sigma_\alpha\Phi^{-1}(\xi_k), \tag{A.2}$$

$$\sigma^2(\xi_k) \equiv \sigma_\alpha^2(1-\rho^2). \tag{A.3}$$

Note that the mean and variance here are consistent with the finite-sample results in Proposition 2 when k/N converges to ξ_k. The mean depends on the order k (shown in Figure A.1), and its shape is very similar to the finite-sample case (Figure 1a). On the other hand, the variance, $\sigma^2(\xi_k)$, is a constant across all quantiles.

A.2 Estimation of ρ and σ_α.

Two key parameters that characterize the distribution of induced order statistics in Propositions 2 and A.2 are ρ, the correlation between unobserved $\boldsymbol{\alpha}$ and $\mathbf{X}$, and σ_α, the cross-sectional standard deviation of α_i. A special case of Proposition 6—equal-weighted portfolios—provides a way to estimate these parameters in practice. Consider an equal-weighted portfolio $\mathcal{S}$ defined in (28) with portfolio weights $\omega_i = 1/n_0$. In this case, Proposition 6 implies that the expected value and variance

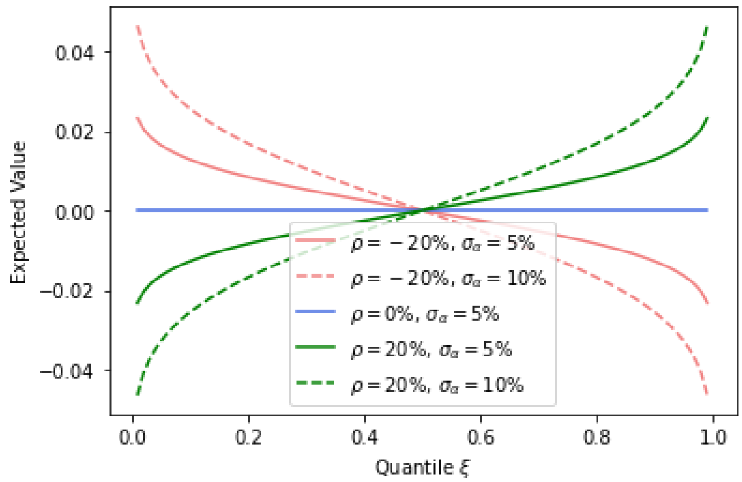


Figure A.1: Asymptotic mean of the induced order statistic, $\alpha_{[i_k:N]}$, as $i_k/N \to \xi \in (0,1)$.

of portfolio alphas are given by:

$$\mathrm{E}\left[\tilde{\alpha}\right] = \frac{\rho\sigma_\alpha}{n_0}\sum_{i\in\mathcal{S}}\mathrm{E}\left[X_{i:N}\right], \tag{A.4}$$

$$\mathrm{Var}\left(\tilde{\alpha}\right) = \sigma_\alpha^2\left(1-\rho^2+\frac{\rho^2}{n_0^2}\left(\sum_{i\in\mathcal{S}}\mathrm{Var}\left(X_{i:N}\right)+2\sum_{i<j\in\mathcal{S}}\mathrm{Cov}\left(X_{i:N},X_{j:N}\right)\right)\right). \tag{A.5}$$

Empirical studies usually report excess returns from equal-weighted portfolios formed by ranking some stock characteristics such as the P/E ratio, book-to-value, or ESG score. As a result, the expected value and variance of the impact-portfolio alpha in (A.4)–(A.5) lead to a natural estimator of these two parameters based on historical data.

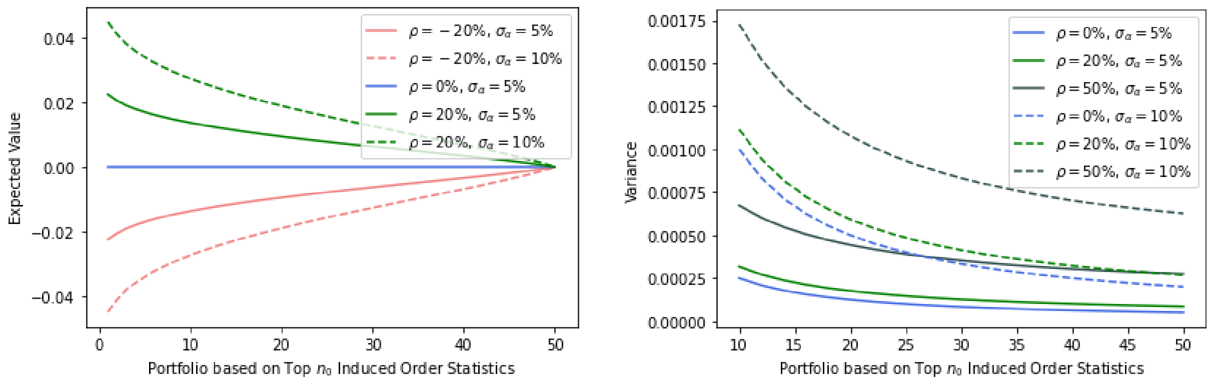
In particular, suppose one empirically measures the portfolio alpha and its variance, which can be substituted into (A.4)–(A.5) to yield a system of two equations with respect to ρ and σ_α, where parameters such as the number of securities in the portfolio (n_0) and the total number of securities in the universe (N) can be easily obtained. This leads, in principle, to a solution for ρ and σ_α.

On the other hand, if the variance in (A.5) is difficult to estimate empirically, one can still use (A.4) to calibrate $\rho\sigma_\alpha$, from which ρ can be solved based on assumptions or prior empirical estimates (such as Pástor and Stambaugh (1999)) for the spread in cross-sectional $\boldsymbol{\alpha}$.

In addition, it is worth emphasizing that the estimation of ρ depends implicitly on the frequency of historical data used to estimate impact-portfolio excess returns, $\tilde{\alpha}$. In theory, if the two terms in (A.4), $\tilde{\alpha}$ and σ_α, both scale linearly as the frequency varies, the estimates of ρ should stay invariant with respect to weekly, monthly, or annual returns. However, they may lead to different empirical estimates in practice, and therefore, the correlation estimated from this procedure should be interpreted in the same frequency space as the return data used.

A.3 Numerical Examples for the Performance of Impact Portfolios

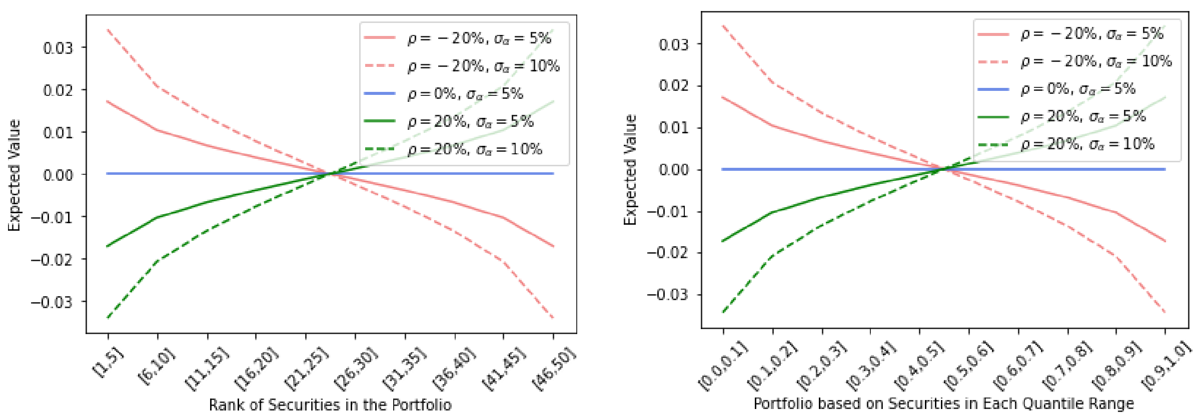
To develop intuition for Proposition 6, consider a portfolio formed by selecting the top n_0 securities based on $\mathbf{X}$. For a market with $N = 50$ securities, Figure A.2 displays the mean and variance of the excess return of portfolios formed in this way. As the number of securities in the portfolio, n_0, increases, the excess return decreases because more securities with weaker alphas are included. At the same time, the variance of the portfolio also decreases thanks to the diversification from more securities.



(a) Expected Value

(b) Variance

Figure A.2: Distribution of portfolio excess return formed by the top n_0 securities ranked by the impact factor, $\mathbf{X}$. The number of total securities, N, is set to be 50.



(a) $N = 50$

(b) $N \to \infty$

Figure A.3: Expected excess return for decile portfolios formed by ranking the impact factor, $\mathbf{X}$. In (a) the number of total securities, N, is set to be 50, and in (b) we show the case when N increases without bound.

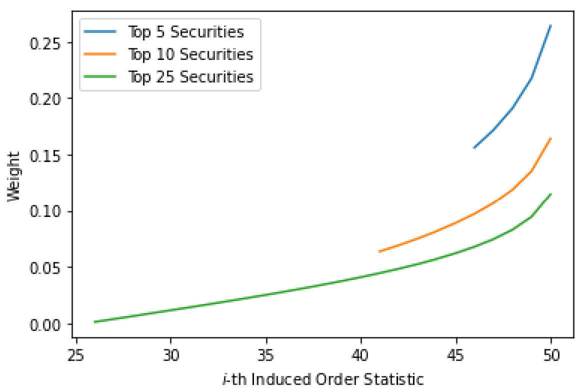
Another typical way of forming portfolios is to sort all securities in the universe into 10 deciles based on $\mathbf{X}$. Figure A.3a contains the expected excess returns of the 10 deciles, which has a similar shape to the expected excess returns of individual securities in Figure 1a.

Finally, we can also consider portfolios as N increases without bound. Suppose we divide the $[0, 1]$ interval into L segments each of length $1/L$, and pick M equally-spaced quantiles within each segment. Specifically, the l-th portfolio is formed by selecting the following quantiles:

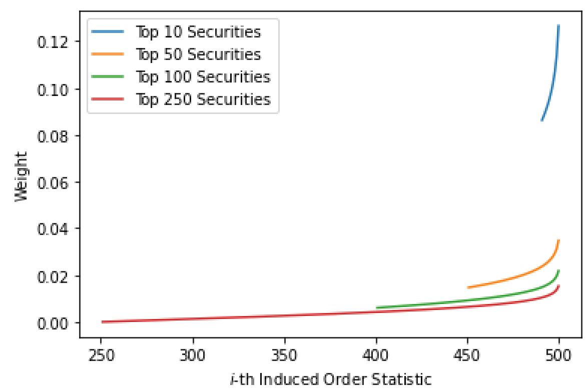
$$\xi_{l,m} = \frac{l - 1 + \frac{m}{M+1}}{L}, \quad m = 1, 2, \cdots, M \tag{A.6}$$

for $l = 1, 2, \cdots, L$. Figure A.3b shows the expected excess returns of this portfolio when $L = M = 10$, which, not surprisingly, has a similar shape to Figure A.3a because the portfolio formed by (A.6) is the limit of the decile portfolio when N increases without bound.

For an illustrative example of Proposition 7, consider a portfolio formed by the top n_0 securities ranked by $\mathbf{X}$, and let n_0 vary from 1 to 250. We assume for a moment that the idiosyncratic volatility is 15% for all securities. Figure A.4 depicts the weights of this portfolio. As expected, securities that rank higher have higher weights. Based on Proposition 7, the weights in Figure A.4 are determined only by the relative rank of the i-th security in the universe of N securities. In other words, changing the correlation, ρ, between $\boldsymbol{\alpha}$ and $\mathbf{X}$ does not affect these weights.



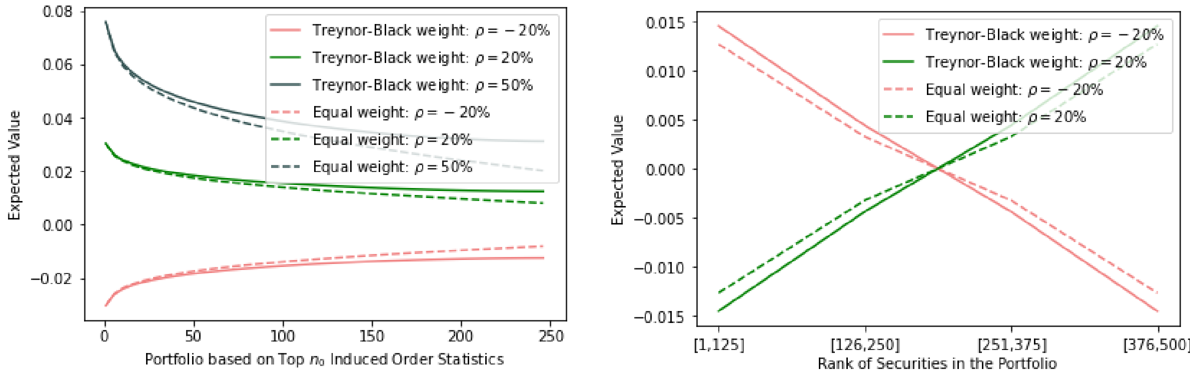
(a) $N = 50$



(b) $N = 500$

Figure A.4: Treynor-Black weights of the securities in the impact portfolio formed by top-ranking securities based on the impact factor, $\mathbf{X}$, with (a) $N = 50$; and (b) $N = 500$.

To further demonstrate the performance of Treynor-Black portfolios as given in Proposition 7, Figure A.5 contains the expected excess return, α_A, for two examples of the impact portfolio in a collection of $N = 500$ securities. Figure A.5a depicts portfolios formed by selecting the top n_0 securities ranked by $\mathbf{X}$. The expected value decreases as n_0 increases and more securities are included. Figure A.5b depicts portfolios formed by dividing all securities into four quantiles based on the ordering of $\mathbf{X}$. In both cases, Treynor-Black portfolios (solid line) achieve higher expected excess returns than the equal-weighted portfolios (dashed line).



(a) Top n_0 Portfolio

(b) Quantile Portfolio

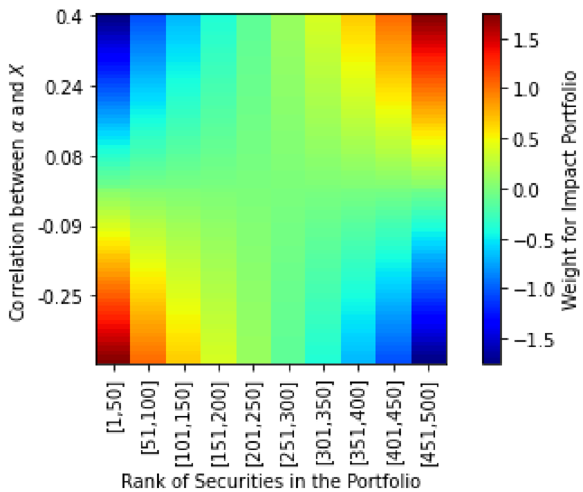
Figure A.5: Expected excess return of the impact portfolio formed based on Treynor-Black weights, with $N = 500$ and $\sigma_\alpha = 5\%$. The expected excess returns of the corresponding equal-weighted portfolios are shown in dashed lines for comparison. (a) shows the case where the top-ranking securities are selected. (b) shows the case where all securities are divided into four segments based on ranking.

A.4 Numerical Examples for Impact Portfolios Combined with Passive Portfolios

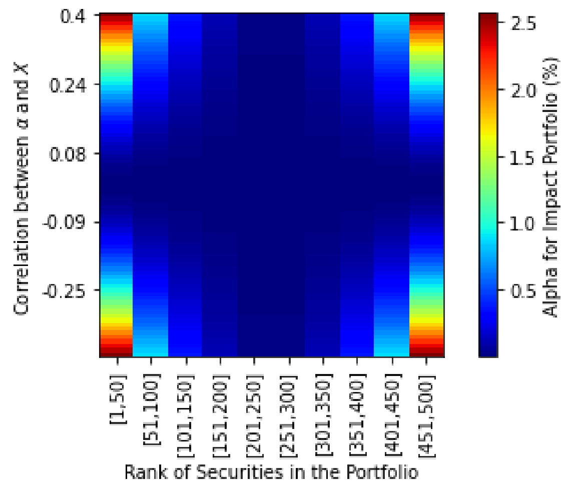
Figure A.6 generalizes the long/short portfolios in Panel B of Table 1, and displays two metrics for the combined portfolio that consists of the impact and passive portfolios, with two different levels of σ_α. In Figures A.6a and A.6b we consider $\sigma_\alpha = 2\%$. In other words, most of the securities have an alpha within $[-4\%, 4\%]$. The weights of the active portfolio range from -1.5 to 1.5, depending on the correlation between $\boldsymbol{\alpha}$ and $\mathbf{X}$ (Figure A.6a). The expected excess return of the combined portfolio ranges from 0% to over 2.5% (Figure A.6b).

In Figures A.6c and A.6d we consider $\sigma_\alpha = 5\%$. In other words, most of the securities have an alpha within $[-10\%, 10\%]$. This is not unimaginable in some highly volatile sectors such as biotech. The weights of the active portfolio can be as high as two, indicating a leveraged impact portfolio and a short position in the passive market portfolio (Figure A.6c). In this case, the expected excess return of the impact portfolio can yield up to 14% (Figure A.6d)!

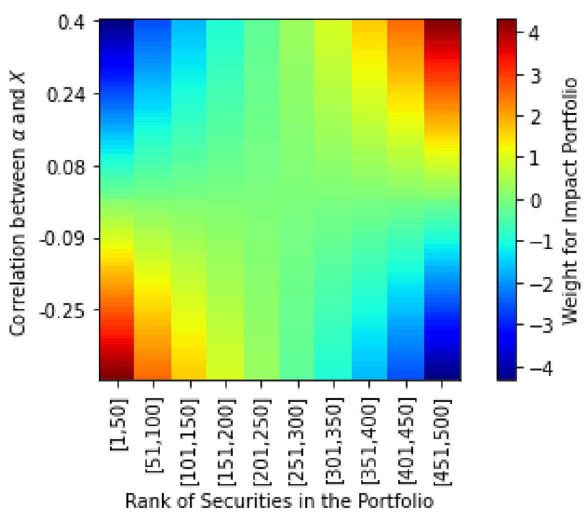
Figure A.7 is the long-only counterpart of Figure A.6, which generalizes the long-only portfolios in the bottom half of Panel B in Table 1. The top-right regions in Figures A.7a and A.7c show positive weights for the impact portfolio, because top-ranking securities yield positive expected excess returns when the correlation, ρ, is positive. Similarly, the bottom-left regions in Figures A.7a and A.7c also show positive weights because these securities, despite having low impact, have positive expected excess returns. As a result, the positive excess returns in the top-right regions in Figures A.7b and A.7d provide a measure of the financial benefit of forming long-only impact portfolios, while the bottom-left regions can be interpreted as the financial cost of not investing in the lowest impact deciles.



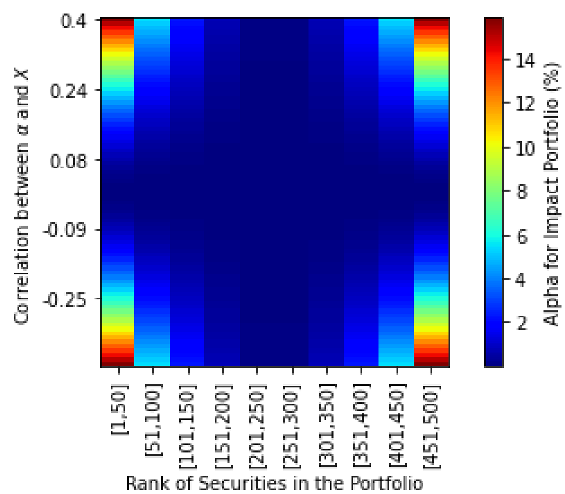
(a) $\sigma_\alpha = 2\%$: Weights



(b) $\sigma_\alpha = 2\%$: Excess Return

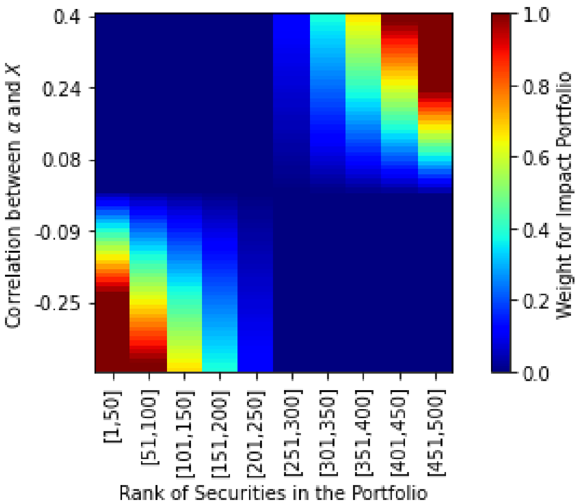


(c) $\sigma_\alpha = 5\%$: Weights

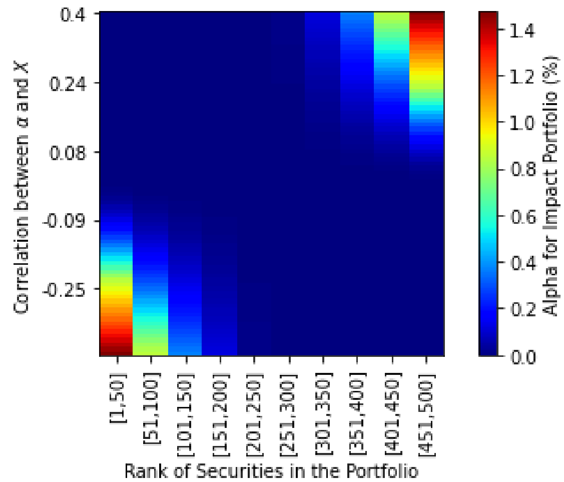


(d) $\sigma_\alpha = 5\%$: Excess Return

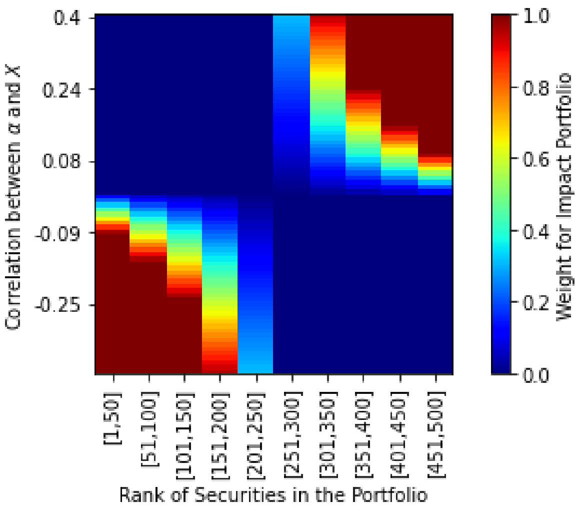
Figure A.6: Performance metrics for the combined long/short portfolio that consists of the impact portfolio with $N = 500$ and passive market portfolio with an annualized risk premium of $\mathrm{E}[R_m] - R_f = 6\%$ and volatility of $\sigma_m = 15\%$. The idiosyncratic volatility is a constant $\sigma(\epsilon_i) = 15\%$ for all securities. (a) and (b) show the Treynor-Black weight for the impact portfolio and the overall expected excess return, respectively, for $\sigma_\alpha = 2\%$. (c) and (d) show the same metrics for $\sigma_\alpha = 5\%$.



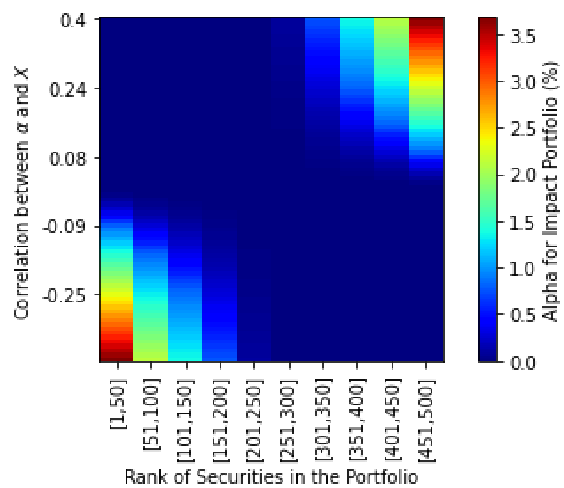
(a) $\sigma_\alpha = 2\%$: Weights



(b) $\sigma_\alpha = 2\%$: Excess Return



(c) $\sigma_\alpha = 5\%$: Weights



(d) $\sigma_\alpha = 5\%$: Excess Return

Figure A.7: Performance metrics for the combined long-only portfolio that consists of the impact portfolio with $N = 500$ and passive market portfolio with an annualized risk premium of $\mathrm{E}[R_m] - R_f = 6\%$ and volatility of $\sigma_m = 15\%$. The idiosyncratic volatility is a constant $\sigma(\epsilon_i) = 15\%$ for all securities. (a) and (b) show the Treynor-Black weight for the impact portfolio and the overall expected excess return, respectively, for $\sigma_\alpha = 2\%$. (c) and (d) show the same metrics for $\sigma_\alpha = 5\%$.

A.5 Institutional Background for the Cystic Fibrosis Foundation.

The CF Foundation is the world's leading philanthropic organization for CF, a rare genetic disease that currently affects more than thirty thousand Americans. Over a period of 12 years, the CF Foundation invested \$150 million to fund CF drug development efforts at Vertex Pharmaceuticals, a Boston-based biotechnology firm. This work led to the identification and development of Kalydeco, the first FDA-approved treatment to address the underlying causes of CF. The Foundation's investment entitled them to receive royalties calculated as a percentage of future sales of successful CF drugs. In 2014, their rights to Vertex royalties were sold to an outside investment firm, New York City-based Royalty Pharma, for \$3.3 billion in cash.

From the financial perspective, a \$3.3 billion return from a \$150 million investment is the dream scenario for any investor, but it could seem like just one individual success story. If we consider CF Foundation's entire portfolio of VP efforts, they allocated a medical and research budget of \$87 million across more than 500 awards in 2012, and over \$160 million across more than 1,100 awards in 2016 (Kim and Lo, 2019). Apparently, from the portfolio perspective, the \$3.3 billion return is still very attractive after factoring in CF Foundation's investments in other projects, even assuming everything else did not produce any financial reward.[1]

In fact, a key part of the CF Foundation's VP strategy has been to divest any ties to commercial products and direct the proceeds to the Foundation's mission as quickly as possible. In contrast to certain types of investment funds, it is not a priority for the CF Foundation to achieve financial returns. Their single purpose is to support CF patients and ease their burden of disease.

The CF Foundation is only one example of VP in biomedicine. More generally, most early-stage drug development programs have low probabilities of success, long time horizons, and large capital requirements (Fagnan et al., 2013), making them less attractive investments than alternatives in other industries like software, social media, telecommunications, etc. In recent years, new tools have emerged to quantify and diversify the risk in these investments (Fagnan et al., 2013; Thakor et al., 2017). Our impact framework provides a systematic approach for constructing impact portfolios and measuring their financial performance, and properly measuring and managing the risk of these investments is the first step towards encouraging more capital to be allocated to accelerate drug development and build greater social value.

Finally, Figure A.8 displays the 250-day rolling-window daily estimated beta of Vertex from 17 July 1992 to 30 December 2020, which we use to estimate the beta used in the CAPM model (43) in the main text.

[1] In fact, since 2014, the CF Foundation has sold additional royalty interests, bringing their total investment returns to over \$4 billion since inception. However, for our purposes, we focus only on the single sale to Royalty Pharma for simplicity since it occurred at a single point in time.

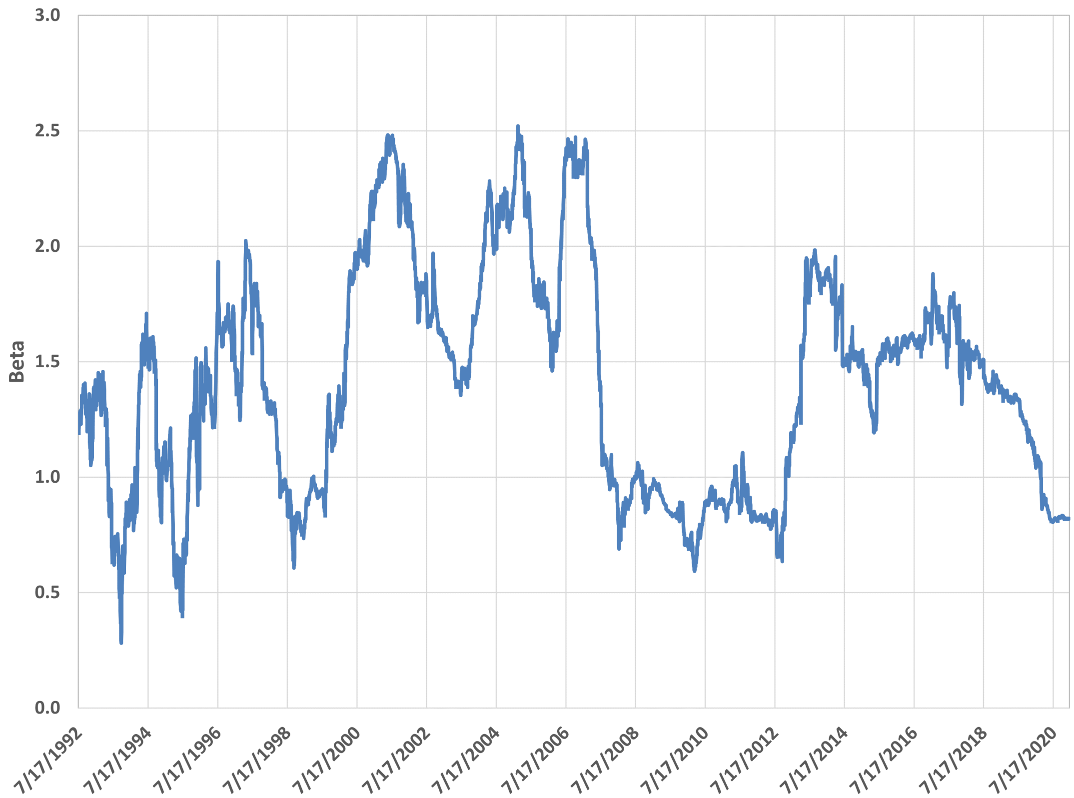


Figure A.8: 250-day rolling-window estimated daily beta coefficients for Vertex Pharmaceuticals from 17 July 1992 to 30 December 2020.

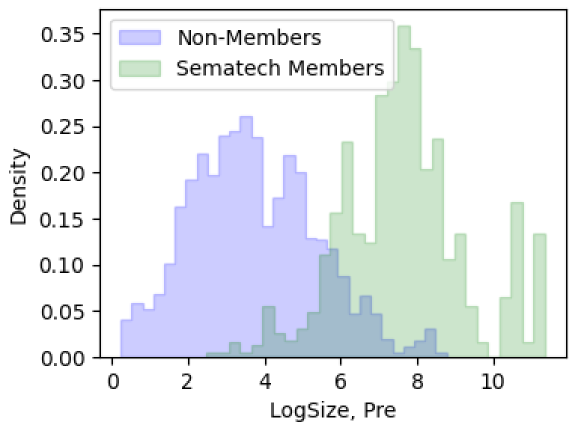
A.6 Endogeneity and Matched Sample for Sematech.

We conduct additional analysis for the Sematech example, in order to assess whether the impact identified in Section 6.2 of the main text can be endogenously explained by other company characteristics.

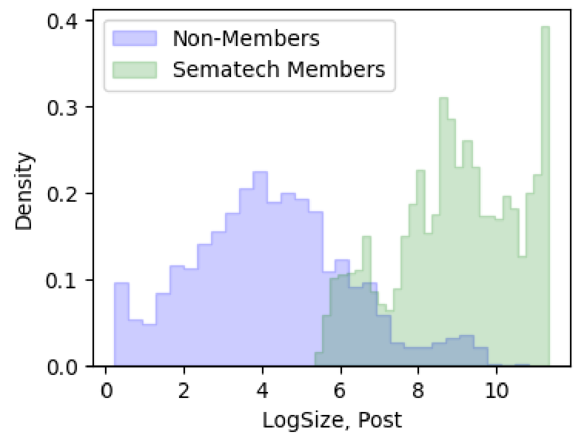
We first compare the logarithm of market capitalization (LogSize) and two profitability measures, including return on asset (RoA) and return on sales (RoS), for Sematech vs. non-Sematech firms before Sematech was formed. Figure A.9 shows the distributions of these firm characteristics computed at the firm-month level before and after Sematech was formed.[2] It is clear that LogSize is the main difference between Sematech and non-Sematech firms. Differences in the two profitability measures also exist before Sematech's formation but to a much lesser extent. The distribution of both profitability measures for Sematech firms shifted towards the right-hand side in the post period, an observation consistent with Irwin and Klenow's (1996a) findings.

We then construct re-weighted samples for non-Sematech firms to match the distribution of Sematech firms in terms of the LogSize, RoA, and RoS, respectively, and show that the return impact from Sematech largely remains robust on these balanced samples.

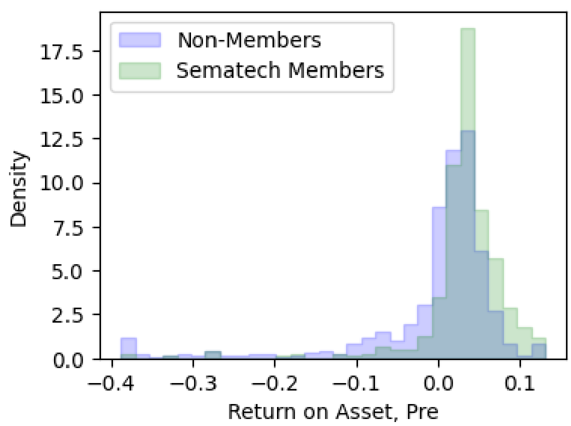
[2]We obtain data from Compustat. The logarithm of market capitalization is winsorized at 1% each side and the two profitability measures are winsorized at 2% each side to reduce the impact of outliers.



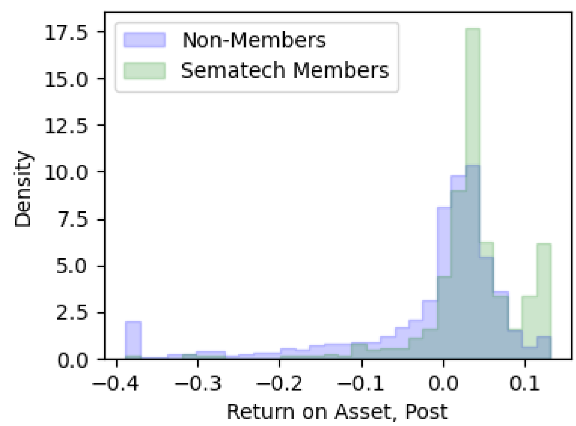
(a) Log Size, Pre



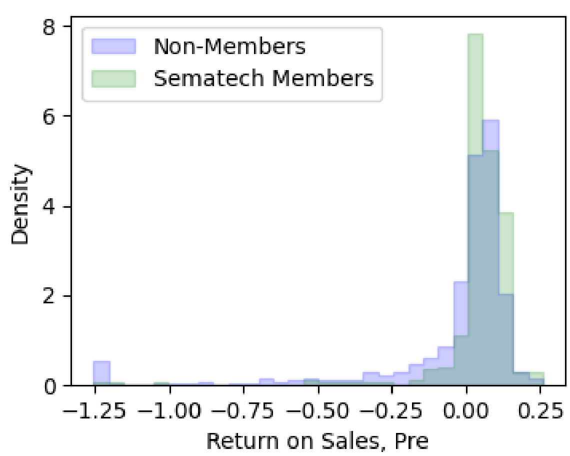
(b) Log Size, Post



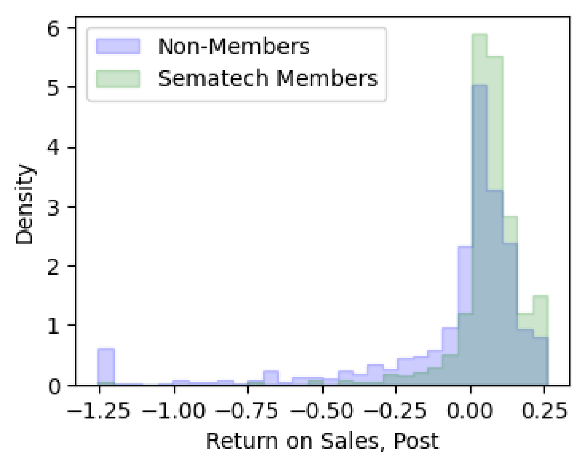
(c) Return on Asset, Pre



(d) Return on Asset, Post



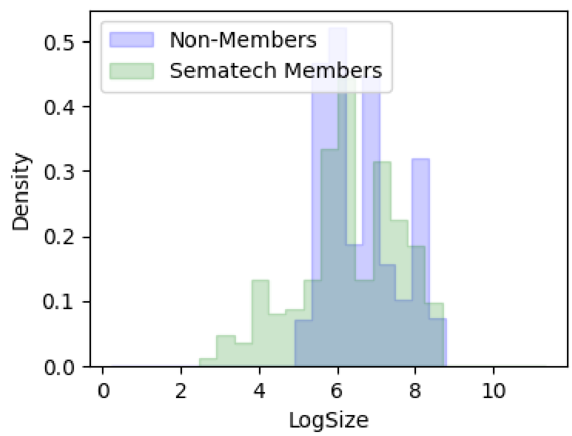
(e) Return on Sales, Pre



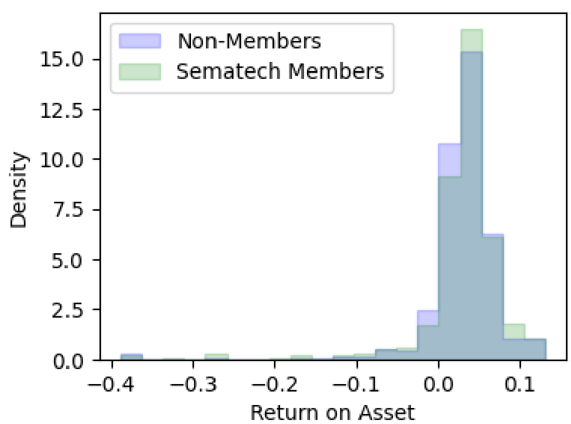
(f) Return on Sales, Post

Figure A.9: Distribution of firm characteristics for Sematech vs non-Sematech firms, before (pre) and after (post) Sematech was formed.

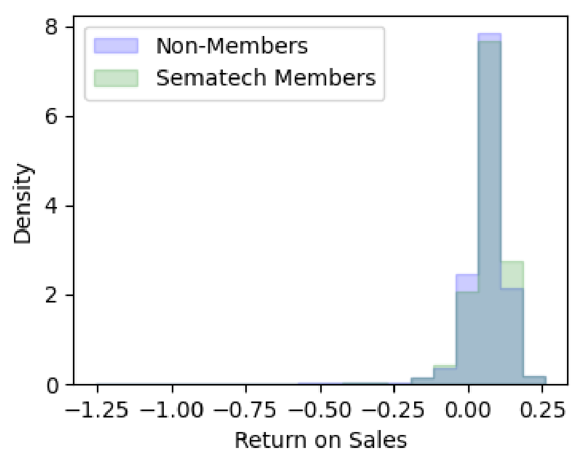
For each firm characteristic (LogSize, RoA, and RoS), we use the average value of each firm in the pre period to estimate two distributions for Sematech and non-Sematech firms. We then assign weights to non-Sematech firms in order to match their pre-period distribution to that of Sematech firms. In addition, a Sematech firm is also assigned zero weight if no non-Sematech firm exists to match its characteristic. Overall, this leads to a small number of firms with nonzero weights in the case of LogSize (10 firms), but a bigger set in the case of RoA (37 firms) and RoS (40 firms). Figure A.10 shows the distributions of weighted samples at the firm-month level, which are much more closely matched between Sematech and non-Sematech firms compared to the original distributions in Figure A.9.



(a) Log Size



(b) Return on Asset



(c) Return on Sales

Figure A.10: Distribution of firm characteristics for Sematech vs non-Sematech firms before Sematech was formed, in which each firm is re-weighted to match the distribution of a corresponding characteristic between Sematech and non-Sematech firms.

Using this weighted sample, Table A.1 summarizes the estimated annualized alphas. Overall, results are consistent with those in Table 2 of the main text, which yields a difference-in-difference estimate of 10.22% lift in annualized excess returns for a firm that joined the R&D consortium.

When samples are weighted by Log Size, the annualized excess return becomes 24.81% though with very wide confidence intervals due to the small number of matched samples. When samples are weighted by profitability measures, the annualized excess returns actually increase to 15.10% for RoA and 12.46% for RoS, both with statistical significance at 5 percent level. Overall, these results suggest that the positive estimate of the financial impact of the Sematech impact factor is likely robust against endogeneity concerns for at least firm size and profitability.

Table A.1: Estimated annualized CAPM excess returns on *weighted* sample, based on a difference-in-difference (DiD) approach that compares Sematech member firms against non-members, before and after the formation of Sematech. Estimated coefficients are significant at the 1 percent (***), 5 percent (**), or 10 percent (*) levels based on Bootstrap confidence intervals.

Firm	Period	Excess Return	Excess Return (Post minus Pre)	Excess Return (DiD)
Panel A: Log Size as Weights				
Sematech Members	Pre	-8.92%	22.13%**	24.81%
	Post	13.20%***		
Non-Members	Pre	-2.54%	-2.68%	
	Post	-5.23%		
Panel B: Return on Asset as Weights				
Sematech Members	Pre	-3.81%	11.01%**	15.10%**
	Post	7.20%***		
Non-Members	Pre	-0.54%	-4.09%	
	Post	-4.63%		
Panel C: Return on Sales as Weights				
Sematech Members	Pre	0.21%	6.23%*	12.46%**
	Post	6.44%**		
Non-Members	Pre	1.61%	-6.22%	
	Post	-4.61%		

In fact, previous research has studied the effects of Sematech on members' R&D spending, profitability, investment, and productivity (Irwin and Klenow, 1996a,b; Link, Teece, and Finan, 1996). In particular, it has been estimated that Sematech induced members to cut their overall R&D spending on the order of $300 million per year, and raised members' profitability relative to non-members' (Irwin and Klenow, 1996a). This reflects more sharing and less duplication of research. In other words, more research has been accomplished per R&D dollar. This provides a potential channel through which excess returns are earned.

A.7 A Simple Execution Model for Meme Stock Trading

Following Bertsimas and Lo's (1998) framework and notation, we assume that an investor seeks to purchase a total of $\bar{S}$ shares of a particular security over a fixed time interval, $[0, T]$. The investor decides how to divide $\bar{S}$ into smaller purchases distributed throughout the interval so as to maximize the final price-impact of the security.[3] The answer depends, of course, on the degree to which a single purchase affects the market price, i.e., the "price impact" and the dynamics of future market prices. Given a particular price-impact function and a specification for the price dynamics, an optimal trading strategy that maximizes the price impact of acquiring $\bar{S}$ in $[0, T]$ may be obtained.

Specifically, denote by S_t the number of shares acquired in period t at price P_t, where $t = 1, 2, \cdots, T$. Then the investor's objective of maximizing final price impact is given by:

$$\max_{\{S_t\}} \mathrm{E}[P_T] \tag{A.7}$$

subject to the constraint that the desired number of shares are acquired:

$$\sum_{t=1}^{T} S_t = \bar{S}. \tag{A.8}$$

We assume that the security price follows the bivariate stochastic process:

$$\begin{aligned} P_t &= P_{t-1} + \theta S_t^z + \gamma F_t + \epsilon_t, \quad \theta > 0,\ z \in (0, 1] \\ F_t &= \delta F_{t-1} + \eta_t, \quad \delta \in (-1, 1) \end{aligned} \tag{A.9}$$

where ϵ_t and η_t are independent white noise processes with mean 0 and variance σ_ϵ^2 and σ_η^2, respectively. The parameter θ specifies the magnitude of the price impact, which is assumed to follow a power law in S_t, where the parameter z specifies the "price sensitivity" of the security or, equivalently, the security's degree of illiquidity. The latter interpretation is motivated by Kyle's (1985) market microstructure model in which liquidity is measured by a loglinear-regression estimate of the log-volume required to move the price by one dollar. Sometimes referred to as "Kyle's lambda," this measure is an inverse proxy of liquidity, with higher values of lambda implying lower liquidity and lower market depth.[4]

The presence of F_t in the law of motion for P_t captures the potential impact of market conditions or private information about the security. For example, F_t can represent new business opportunities created by the company, the impact of popular sentiment—as in the case of GME, as well as any of the other factors mentioned above. In either case, the impact of F_t on trading profits and the time series properties of F_t both have important implications for the feasibility and profitability of price-impact investing. With these price dynamics, the following result completely characterizes the optimal price-impact strategy and its corresponding expected profit:

Proposition A.3. *Under the price dynamics specified by (A.9), the strategy that maximizes the*

[3] Note that this is *not* the objective function considered by Bertsimas and Lo (1998)—the problem they pose is how to divide $\bar{S}$ so as to maximize *cumulative profits*, which they solve via stochastic dynamic programming.

[4] See also Lillo, Farmer, and Mantegna (2003) and Almgren et al. (2005) for more detailed explorations of the power law of price impact in equity markets. When $z=1$, this reduces to the "linear price impact with information" specification from Bertsimas and Lo (1998).

total price impact, (A.7), is given by:

$$S_1 = S_2 = \cdots = S_T = \frac{\bar{S}}{T}, \tag{A.10}$$

and its corresponding expected profit is given by:

$$V^* = \left(\frac{\theta \bar{S}^z (T-1)}{2T^z} + \frac{\gamma \delta F_1 \left(1 - T\delta^{T-1} + (T-1)\delta^T\right)}{(1-\delta)^2 T} \right) \bar{S}. \tag{A.11}$$

In fact, when $z = 1$ and price impact is linear in trading quantities,[5] it does not matter how trades are allocated because the total impact from T trades is always equal to the impact of one single trade of size $\bar{S}$. However, when the price impact is a concave function in general ($0 < z < 1$), the optimal strategy is to simply divide the total order $\bar{S}$ into T equal "waves," and trade them at regular intervals, as specified in (A.10).

The expression for V^* in (A.11) shows that the expected profit of price-impact investing depends on two factors: the market impact as parameterized by θ and z, and influences from other factors (sentiment, liquidity, private information, etc.) as parameterized by γ and the AR(1) coefficient governing these other factors (δ).

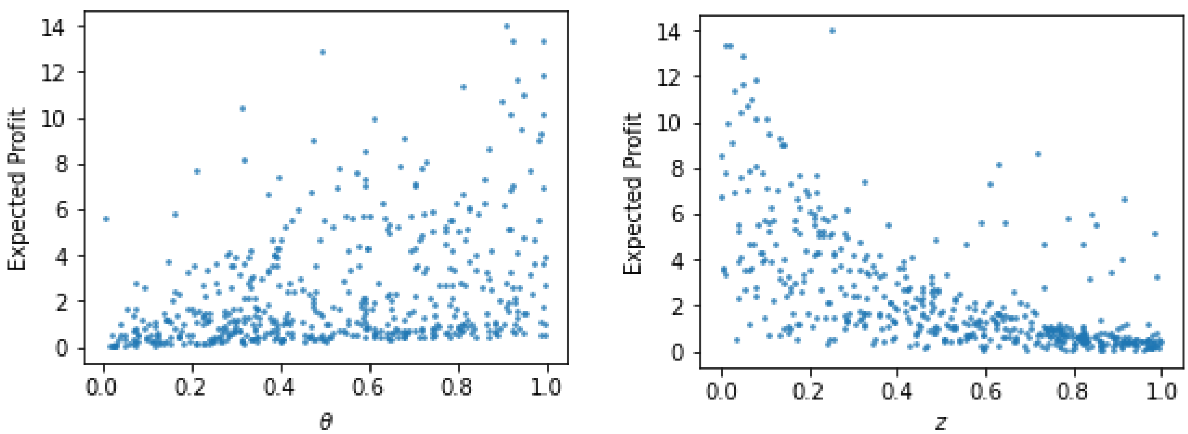
To illustrate the effect of these parameters on trading profit V^*, we simulate a universe of $N = 500$ securities where the parameters, θ, z, γ, and δ, are generated by four independent uniform distributions on $[0, 1]$. In the following analysis, we assume that the first realization of $X_1 = 1$, without loss of generality.

In Figure A.11a, we first show the relationship of the expected profit V^* with respect to market impact (θ). As θ increases, expected profit increases as well. This is quite intuitive because the stronger the market impact, the easier it is for short squeezers to induce price momentum and generate profits. If we consider a collection of securities each with a different θ, the correlation between their market-impact coefficients and expected profit is 37%, implying that sorting $\boldsymbol{\alpha}$ based on θ will generate positive excess returns in the context of our impact-investing framework.

Figure A.11b displays the relation between expected profit V^* and sensitivity z. As power increases from 0 to 1, the expected profit decreases. This is because lower values of z correspond to more concave price-impact functions, for which each small trading segment has larger price impact. The correlation between z and expected profit is -63%. In other words, one can achieve positive excess returns by selecting securities based on the reverse ordering of sensitivity z.

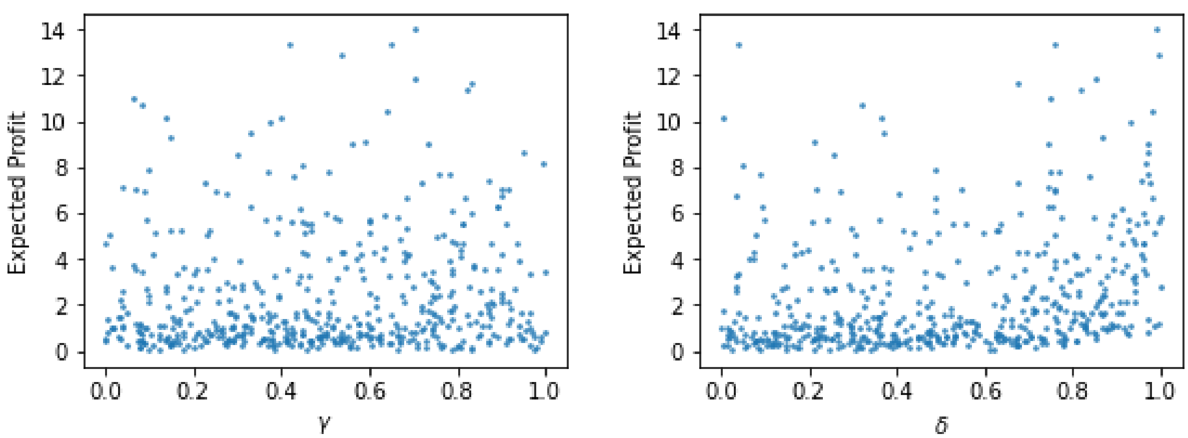
Figure A.11c displays the relation between expected profit V^* and influences from other factors (γ), which has a weak positive correlation of 9%. Finally, Figure A.11d displays the relation between AR coefficient (δ) and expected profit. The expected profit is larger when δ is larger. This is because we have assumed the first realization of F_t is positive, and higher autocorrelations imply stronger momentum. Indeed, the correlation between the AR coefficient, δ, and expected profit is 26% in this simulated market.

[5] See also Bertsimas and Lo (1998).



(a) Price impact (θ)

(b) Market sensitivity (z)



(c) Influences from other factors (γ)

(d) AR coefficient (δ)

Figure A.11: The expected profit, V^*, of price-impact investing as a function of four parameters in (A.9), for a market with $N=500$ securities with simulated parameters. Here we set $\theta=1$, $z=1$, $\gamma=1$, $\delta=10\%$, $\bar{S}=1$, $F_1=1$, and $T=30$ by default, and vary each parameter accordingly.

We summarize the results from Figure A.11 in Table A.2, and provide their implied α when applied to a collection of 500 securities simultaneously, each with different price dynamics as specified in (A.9). Panels A, B, and C show the expected excess returns if investors apply θ, z, γ, and δ, respectively, to rank securities, where the correlations with trading profits are obtained from our simple execution model. The expected α can be very high with a leveraged portfolio, driven by the high correlation between stock $\boldsymbol{\alpha}$ and the price-impact investing factor in certain cases.

In practice, it is difficult to accurately calibrate the relevant parameters for each stock, hence the expected profit of engaging in GME-like price-impact investing is correspondingly difficult to estimate. However, this example highlights the fundamental determinants for a price-impact investor's α: the correlation between each stock's trading profit and stock characteristics, e.g., market capitalization, liquidity, specific forms of market impact, attention from the general public, main shareholders, short interest, or anything correlated with stock returns. Higher correlations lead to higher alpha when following that particular characteristic to select target stock.

Table A.2: Estimated excess returns per annum for the price-impact investing factor, based on the optimal strategy's profit in (A.11) and its implied correlations with respect to various characteristics of individual securities. Here we assume $\sigma_\alpha = 5\%$—an intermediate value based on Pástor and Stambaugh's (1999) estimated range of σ_α (between 0% to 10%), and the passive portfolio has an annualized risk premium of $\mathrm{E}[R_m] - R_f = 6\%$ and volatility of $\sigma_m = 15\%$.

Impact Portfolio	Weight of Impact Portfolio	Expected Excess Return	
		Impact Portfolio	Combined with Passive Portfolio
	ω_A	α_A	$\omega_A\alpha_A$
Panel A: Ranking based on price impact (θ)			
Model-implied correlation with alpha: $\rho = 37\%$.			
Top Half	9.19	2.3%	21.4%
Top Decile	4.09	3.4%	14.0%
Top 2%	1.20	4.5%	5.4%
Panel B: Ranking based on market sensitivity (z); reverse order			
Model-implied correlation with alpha: $\rho = 63\%$.			
Top Half	15.58	3.9%	61.5%
Top Decile	6.95	5.8%	40.3%
Top 2%	2.04	7.6%	15.5%
Panel C: Ranking based on other factors (γ)			
Model-implied correlation with alpha: $\rho = 9\%$.			
Top Half	2.20	0.6%	1.2%
Top Decile	0.98	0.8%	0.8%
Top 2%	0.29	1.1%	0.3%
Panel D: Ranking based on AR coefficient for other factors (δ)			
Model-implied correlation with alpha: $\rho = 26\%$.			
Top Half	6.43	1.6%	10.5%
Top Decile	2.87	2.4%	6.9%
Top 2%	0.84	3.2%	2.7%

B Proofs

In this Appendix, we provide proofs for all the propositions.

B.1 Proof of Proposition 1

The constraints on the right-hand side optimization problem of (3) is a subset of the left-hand side optimization problem. Therefore the inequality follows.

To give a bound on the utility loss between the unconstrained portfolio W and the constrained portfolio W^c, we consider an intermediate portfolio W^{c1} that is also constrained to the subset $\mathcal{S}$, but with equal factor loadings as the unconstrained portfolio W. In other words, the portfolio weights for W^{c1} satisfy the following conditions:

$$\omega_i^{c1} = 0 \text{ for } i \notin \mathcal{S} \tag{A.12}$$

$$\sum_{i \in \mathcal{S}} \omega_i^{c1} = 1 \tag{A.13}$$

$$\sum_{i=1}^{N} \omega_i \beta_{ik} = \sum_{i \in \mathcal{S}} \omega_i^{c1} \beta_{ik} \text{ for } k = 1, \ldots, K. \tag{A.14}$$

Because W^c maximizes the utility in (3),

$$\begin{aligned}
\mathrm{E}[U(W^c)] &= \mathrm{E}\left[U\left(R_f + \sum_{k=1}^{K}\sum_{i \in \mathcal{S}} \omega_i^c \beta_{ik}\left(\Lambda_k - R_f\right) + \sum_{i \in \mathcal{S}} \omega_i^c \epsilon_i\right)\right] \\
&\geq \mathrm{E}\left[U\left(R_f + \sum_{k=1}^{K}\sum_{i \in \mathcal{S}} \omega_i^{c1} \beta_{ik}\left(\Lambda_k - R_f\right) + \sum_{i \in \mathcal{S}} \omega_i^{c1} \epsilon_i\right)\right] \\
&= \mathrm{E}\left[U\left(R_f + \sum_{k=1}^{K}\sum_{i=1}^{N} \omega_i \beta_{ik}\left(\Lambda_k - R_f\right) + \sum_{i \in \mathcal{S}} \omega_i^{c1} \epsilon_i\right)\right] = \mathrm{E}[U(W^{c1})].
\end{aligned} \tag{A.15}$$

Now we consider the utility of the following two portfolios,

$$\begin{aligned}
\mathrm{E}[U(W)] &= \mathrm{E}\left[U\left(R_f + \sum_{k=1}^{K}\sum_{i=1}^{N} \omega_i \beta_{ik}\left(\Lambda_k - R_f\right) + \sum_{i=1}^{N} \omega_i \epsilon_i\right)\right], \\
\mathrm{E}[U(W^{c1})] &= \mathrm{E}\left[U\left(R_f + \sum_{k=1}^{K}\sum_{i=1}^{N} \omega_i \beta_{ik}\left(\Lambda_k - R_f\right) + \sum_{i \in \mathcal{S}} \omega_i^{c1} \epsilon_i\right)\right].
\end{aligned} \tag{A.16}$$

Note that they only differ in the last term in the parenthesis, the idiosyncratic volatilities. Denote $A \equiv R_f + \sum_{k=1}^{K}\sum_{i=1}^{N} \omega_i \beta_{ik}\left(\Lambda_k - R_f\right)$ and $B \equiv \sum_{i=1}^{N} \omega_i \epsilon_i$ (or $\sum_{i \in \mathcal{S}} \omega_i^{c1} \epsilon_i$). For any well-behaved utility function U, because $\mathrm{E}[B]=0$, we have:

$$\mathrm{E}[U(A+B)] \approx \mathrm{E}\left[U(A) + U'(A)B + \frac{1}{2}U''(A)B^2\right] = \mathrm{E}[U(A)] + \frac{1}{2}\mathrm{E}[U''(A)]\mathrm{Var}[B^2] \tag{A.17}$$

by second-order Taylor expansion around $B=0$. Since $\mathrm{E}[U(W)]$ and $\mathrm{E}[U(W^{c1})]$ differs only through the idiosyncratic volatility term, B, we have:

$$\begin{aligned} \mathrm{E}[U(W)] - \mathrm{E}[U(W^c)] &\leq \mathrm{E}[U(W)] - \mathrm{E}[U(W^{c1})] \\ &\approx \frac{1}{2}\mathrm{E}[U''(A)]\left(\mathrm{Var}\left(\sum_{i=1}^{N}\omega_i\epsilon_i\right) - \mathrm{Var}\left(\sum_{i\in\mathcal{S}}\omega_i^{c1}\epsilon_i\right)\right). \end{aligned} \tag{A.18}$$

When the number of securities, N, is large, suppose further that:

$$\omega_i \approx \frac{1}{N},\ \omega_i^{c1} \approx \frac{1}{N-n},\ \sigma(\epsilon_i) \approx \sigma_\epsilon, \tag{A.19}$$

where n is the number of securities excluded in $\mathcal{S}$, and σ_ϵ is the common idiosyncratic volatility for all securities. We have:

$$\mathrm{E}[U(W)] - \mathrm{E}[U(W^c)] \leq \frac{1}{2}\mathrm{E}[U''(A)]\left(\frac{\sigma_\epsilon^2}{N} - \frac{\sigma_\epsilon^2}{N-n}\right) = -\frac{1}{2}\mathrm{E}[U''(A)]\sigma_\epsilon^2\left(\frac{n}{N(N-n)}\right). \tag{A.20}$$

When the number of securities excluded in $\mathcal{S}$, n, is small relative to the total number of securities, N, the utility loss (A.20) is also small.

Finally, we observe that the assumptions in (A.19) are non-critical for our main conclusions here, and can be relaxed at the expense of simplicity of the mathematical derivation.

B.2 Proof of Proposition 2

Because $\mathbf{X}$ and $\boldsymbol{\alpha}$ are jointly normal, we can express α_i with the following linear relationship:

$$\alpha_i = \mu_\alpha + \rho\frac{\sigma_\alpha}{\sigma_x}(X_i - \mu_x) + e_i, \tag{A.21}$$

where e_i are normal random variables with $\mathrm{E}[e_i] = 0$ and $\mathrm{Var}(e_i) = \sigma_\alpha^2(1-\rho^2)$, and the X_i and the e_i are mutually independent. Ordering securities based on X_i, we have:

$$\alpha_{[i:N]} = \mu_\alpha + \rho\frac{\sigma_\alpha}{\sigma_x}(X_{i:N} - \mu_x) + e_{[i]}, \tag{A.22}$$

where $e_{[i]}$ denotes the particular e_i associated with $X_{i:N}$. Note that $X_{i:N}$ on the right-hand side are the usual order statistics, while $\alpha_{[i:N]}$ on the left-hand side are the *induced* order statistics. Because X_i and e_i are independent, the set of $X_{i:N}$ and the set of $e_{[i]}$ are also independent. Therefore, we can calculate the first two moments of $\boldsymbol{\alpha}$ based on the relationship in (A.22):

$$\mathrm{E}\left[\alpha_{[i:N]}\right] = \mu_\alpha + \rho\frac{\sigma_\alpha}{\sigma_x}(\mathrm{E}[X_{i:N}] - \mu_x) + e_{[i]} = \mu_\alpha + \rho\sigma_\alpha\mathrm{E}[X_{i:N}], \tag{A.23}$$

$$\mathrm{Var}\left(\alpha_{[i:N]}\right) = \rho^2\frac{\sigma_\alpha^2}{\sigma_x^2}\mathrm{Var}\left(X_{i:N}\right) + \sigma_\alpha^2(1-\rho^2) = \sigma_\alpha^2\left(1-\rho^2+\rho^2\mathrm{Var}\left(X_{i:N}\right)\right), \tag{A.24}$$

$$\mathrm{Cov}\left(\alpha_{[i:N]}, \alpha_{[j:N]}\right) = \mathrm{Cov}\left(\rho\frac{\sigma_\alpha}{\sigma_x}X_{i:N}, \rho\frac{\sigma_\alpha}{\sigma_x}X_{j:N}\right) = \sigma_\alpha^2\rho^2\mathrm{Cov}\left(X_{i:N}, X_{j:N}\right). \tag{A.25}$$

See also David and Nagaraja (2004, Section 6.8).

B.3 Proof of Proposition 3

We first observe that $U_{i:N} \equiv \Phi(X_{i:N})$ maps the i-th normal order statistics to the i-th order statistics from a uniform distribution on $[0,1]$, where Φ is the cumulative distribution function of standard normal random variables. We define $Q \equiv \Phi^{-1}$ and write $X_{i:N} = Q(U_{i:N})$. We then expand $Q(U_{i:N})$ in a Taylor series around the expected value of $Q(U_{i:N})$:

$$\mathrm{E}[Q(U_{i:N})] = \frac{i}{n+1} = p_i, \tag{A.26}$$

which gives:

$$X_{i:N} = Q(U_{i:N}) = Q(p_i) + (U_{i:N} - p_i)\,Q'(p_i) + \frac{1}{2}(U_{i:N} - p_i)^2\,Q''(p_i) + \frac{1}{6}(U_{i:N} - p_i)^3\,Q'''(p_i) + \cdots. \tag{A.27}$$

Substituting (A.27) into the definition of $\mathrm{E}\left[X_{i:N}\right]$, $\mathrm{Var}\left(X_{i:N}\right)$, and $\mathrm{Cov}\left(X_{i:N}, X_{j:N}\right)$, and rearranging the terms lead to (11)-(13) in Proposition 3. See also David and Nagaraja (2004, Section 4.6).

In particular, for standard normal random variables we have $Q'(p_i) = 1/\phi(Q)$ where ϕ is the density function for standard normal random variables. Therefore we can calculate:

$$Q''(p_i) = \frac{d\left(1/\phi(Q)\right)}{d\Phi(Q)} = \frac{d\left(1/\phi(Q)\right)}{dQ}\frac{dQ}{d\Phi(Q)} = \frac{Q}{\phi^2(Q)}, \tag{A.28}$$

$$Q'''(p_i) = \frac{1+2Q^2}{\phi^3(Q)}, \tag{A.29}$$

$$Q''''(p_i) = \frac{Q(7+6Q^2)}{\phi^4(Q)}, \tag{A.30}$$

which completes the proof for (14)-(17).

B.4 Proof of Proposition 4

Because $\mathbf{X}$ and $\boldsymbol{\alpha}$ are both normally distributed, we observe that $\frac{X_i-\mu_x}{\sigma_x}$ and $\frac{\alpha_i-\mu_\alpha}{\sigma_\alpha}$ both follow the standard normal distribution. Therefore,

$$\mathrm{E}\left[\frac{X_{i:N}-\mu_x}{\sigma_x}\right] = \mathrm{E}\left[\frac{\alpha_{i:N}-\mu_\alpha}{\sigma_\alpha}\right], \tag{A.31}$$

$$\mathrm{Var}\left(\frac{X_{i:N}-\mu_x}{\sigma_x}\right) = \mathrm{Var}\left(\frac{\alpha_{i:N}-\mu_\alpha}{\sigma_\alpha}\right), \tag{A.32}$$

$$\mathrm{Cov}\left(\frac{X_{i:N}-\mu_x}{\sigma_x}, \frac{X_{j:N}-\mu_x}{\sigma_x}\right) = \mathrm{Var}\left(\frac{\alpha_{i:N}-\mu_\alpha}{\sigma_\alpha}, \frac{\alpha_{j:N}-\mu_\alpha}{\sigma_\alpha}\right). \tag{A.33}$$

We have assumed, without loss of generality, that $\mu_\alpha = \mu_x = 0$ and $\sigma_x = 1$, which leads to:

$$\mathrm{E}[X_{i:N}] = \frac{\mathrm{E}[\alpha_{i:N}]}{\sigma_\alpha}, \tag{A.34}$$

$$\mathrm{Var}\left(X_{i:N}\right) = \frac{\mathrm{Var}\left(\alpha_{i:N}\right)}{\sigma_\alpha^2}, \tag{A.35}$$

$$\mathrm{Cov}\left(X_{i:N}, X_{j:N}\right) = \frac{\mathrm{Cov}\left(\alpha_{i:N}, \alpha_{j:N}\right)}{\sigma_\alpha^2}. \tag{A.36}$$

This together with (8)-(10) gives:

$$\mu_i = \mathrm{E}\left[\alpha_{[i:N]}\right] = \rho\sigma_\alpha \mathrm{E}\left[X_{i:N}\right] = \rho \mathrm{E}[\alpha_{i:N}]. \tag{A.37}$$

$$\sigma_i^2 - \sigma_\alpha^2 = \sigma_\alpha^2\rho^2\left[\mathrm{Var}\left(X_{i:N}\right) - 1\right] = \rho^2\left[\mathrm{Var}\left(\alpha_{i:N}\right) - \sigma_\alpha^2\right], \tag{A.38}$$

$$\sigma_{ij} \equiv \mathrm{Cov}\left(\alpha_{[i:N]}, \alpha_{[j:N]}\right) = \sigma_\alpha^2\rho^2 \mathrm{Cov}\left(X_{i:N}, X_{j:N}\right) = \rho^2 \mathrm{Cov}\left(\alpha_{i:N}, \alpha_{j:N}\right). \tag{A.39}$$

B.5 Proof of Proposition 5

For simplicity, we define $\boldsymbol{\lambda} \equiv [\,\lambda_1 \cdots \lambda_N\,]^T$, and observe that $\mathbf{X}$ and $\boldsymbol{\lambda}$ can be rewritten as:

$$\begin{aligned} \mathbf{X} &= \mu_x \mathbf{1} + \mathbf{C}_x \mathbf{N}_x \\ \boldsymbol{\lambda} &= \mu_\lambda \mathbf{1} + \mathbf{C}_\lambda \mathbf{N}_\lambda \end{aligned} \tag{A.40}$$

where $\mathbf{1} \equiv [\,1 \cdots 1\,]^T$ is a column vector of ones with size N, $\mathbf{N}_x$ and $\mathbf{N}_y$ are both N-dimensional standard normal random vectors with $\mathrm{Cov}\left(\mathbf{N}_x, \mathbf{N}_y\right) = \boldsymbol{\Sigma}$, and $\mathbf{C}_x$ and $\mathbf{C}_y$ are both $N \times N$ deterministic matrices. The specification in (A.40) completely characterizes the joint distribution of $\mathbf{X}$ and $\boldsymbol{\lambda}$. In light of the parameterization in Assumption (A2), we have:

$$\begin{aligned} \mathbf{C}_x &= \sqrt{1-\rho_x}\sigma_x \mathbf{I} + \left(\sqrt{1+(N-1)\rho_x} - \sqrt{1-\rho_x}\right)\sigma_x \mathbf{L} \\ \mathbf{C}_\lambda &= \sqrt{1-\rho_\lambda}\sigma_\lambda \mathbf{I} + \left(\sqrt{1+(N-1)\rho_\lambda} - \sqrt{1-\rho_\lambda}\right)\sigma_\lambda \mathbf{L} \\ \boldsymbol{\Sigma} &= \frac{\rho_{x\lambda} - \tilde{\rho}_{x\lambda}}{\sqrt{(1-\rho_x)(1-\rho_\lambda)}}\mathbf{I} + \left(\frac{\rho_{x\lambda} + (n-1)\tilde{\rho}_{x\lambda}}{\sqrt{(1+(n-1)\rho_x)(1+(n-1)\rho_\lambda)}} - \frac{\rho_{x\lambda} - \tilde{\rho}_{x\lambda}}{\sqrt{(1-\rho_x)(1-\rho_\lambda)}}\right)\mathbf{L} \end{aligned} \tag{A.41}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{L} \equiv \frac{\mathbf{1} \cdot \mathbf{1}^T}{N}$ is a matrix whose elements are all $1/N$.

We now define a projection matrix:

$$\begin{aligned}
\mathbf{P} &\equiv \left(\mathbf{C}_\lambda \mathbf{\Sigma}^T \mathbf{C}_x^T\right)\left(\mathbf{C}_x \mathbf{C}_x^T\right)^{-1} \\
&= \underbrace{\rho_{adj}\frac{\sigma_\lambda}{\sigma_x}}_{a}\mathbf{I} + \underbrace{\left(\frac{\rho_{x\lambda} + (n-1)\tilde{\rho}_{x\lambda}}{1 + (n-1)\rho_x} - \rho_{adj}\right)\frac{\sigma_\lambda}{\sigma_x}}_{b}\mathbf{L} \\
&= a\mathbf{I} + b\mathbf{L}
\end{aligned} \tag{A.42}$$

and it is easy to show that:

$$\boldsymbol{\lambda} - \mathbf{P}\mathbf{X} \perp \mathbf{X}. \tag{A.43}$$

Therefore, when assuming $\mu_x = 0$ and $\sigma_x = 1$, we have:

$$\begin{aligned}
\mathrm{E}\left[\lambda_{[i:N]}\right] &= \mathrm{E}\left[(\boldsymbol{\lambda} - \mathbf{P}\mathbf{X})_{[i:N]}\right] + \mathrm{E}\left[(\mathbf{P}\mathbf{X})_{[i:N]}\right] \\
&= \mu_\lambda - (a+b)\mu_x + a\mathrm{E}\left[X_{i:N}\right] + b\mu_x \\
&= \mu_\lambda + \rho_{adj}\sigma_\lambda \mathrm{E}\left[X_{i:N}\right],
\end{aligned} \tag{A.44}$$

which proves (25). The variances and covariances in (26)-(27) can be proven similarly following the same decomposition in (A.44). See also Lee and Viana (1999).

B.6 Proof of Proposition 6

The expected excess return follows directly from the distribution of alphas for single securities in Proposition 2. The variance also follows by rearranging terms:

$$\begin{aligned}
\mathrm{Var}\left(\tilde{\alpha}\right) =& \sum_{i \in P} \omega_i^2 \sigma_i^2 + 2 \sum_{i<j \in P} \omega_i \omega_j \sigma_{ij} \\
=& \sigma_\alpha^2 \left(1 - \rho^2 + \rho^2 \sum_{i \in P} \omega_i^2 \mathrm{Var}\left(X_{i:N}\right) + 2\rho^2 \sum_{i<j \in P} \omega_i \omega_j \mathrm{Cov}\left(X_{i:N}, X_{j:N}\right)\right) \\
=& \sigma_\alpha^2 \left(1 - \rho^2 + \rho^2 \left(\sum_{i \in P} \omega_i^2 \mathrm{Var}\left(X_{i:N}\right) + 2 \sum_{i<j \in P} \omega_i \omega_j \mathrm{Cov}\left(X_{i:N}, X_{j:N}\right)\right)\right).
\end{aligned} \tag{A.45}$$

B.7 Proof of Proposition 7

Because of the decomposition in (32), and the fact that ζ_i are independent of ϵ_i, the combined idiosyncratic variance for security i is simply $\sigma_i^2 + \sigma(\epsilon_i)^2$, where σ_i^2 is the variance of the i-th induced order statistic given in (9), and $\sigma(\epsilon_i)^2$ is the original idiosyncratic variance for security i given in (1). The classical result of Treynor and Black (1973) maintains that to maximize the Sharpe ratio of the portfolio, security weights should be proportional to the expected excess returns divided by the idiosyncratic variance, which proves (33).

(34) follows from plugging in results from Proposition A.2 into (33).

B.8 Proof of Proposition 8

By definitions in (35)-(36), the return of the impact portfolio in excess of the risk-free rate can be written as:

$$R_A - R_f = \alpha_A + \beta_A(R_m - R_f) + \epsilon_A. \tag{A.46}$$

When combining with the passive market portfolio, the weight of the impact portfolio, ω_A, is given in (38). Therefore, the return of the combined portfolio, in excess of the risk-free rate, is

$$\begin{aligned} R_P - R_f &= \omega_A(R_A - R_f) + (1 - \omega_A)(R_m - R_f) \\ &= \omega_A\left(\alpha_A + \beta_A(R_m - R_f) + \epsilon_A\right) + (1 - \omega_A)(R_m - R_f) \\ &= \omega_A\alpha_A + (R_m - R_f)(\beta_A\omega_A + (1 - \omega_A)) + \omega_A\epsilon_A, \end{aligned} \tag{A.47}$$

which completes the proof of (39). (40) and (41) follow directly from simple calculations of the expected value and variance of R_P based on (A.47).

B.9 Proof of Proposition A.1

This proposition follows from Yang (1977). See also Lo and MacKinlay (1990) for an application in a different context.

B.10 Proof of Proposition A.2

This follows from Proposition A.1 by observing that $\Phi(\xi_k) = F_x(\xi_k\sigma_x + \mu_x)$. Alternatively, this result can be proved by taking the limit as $N \to \infty$ based on the finite-sample results in Proposition 2-3.

B.11 Proof of Proposition A.3

Based on the price process, (A.9), the investor's objective, (A.7), can be written as:

$$\begin{aligned} \mathrm{E}[P_T] &= \mathrm{E}[P_{T-1} + \theta S_T^z + \gamma F_T + \epsilon_T] \\ &= \mathrm{E}[P_{T-2} + \theta S_{T-1}^z + \gamma F_{T-1} + \epsilon_{T-1} + \theta S_T^z + \gamma F_T + \epsilon_T] \\ &= P_0 + \theta(S_1^z + \cdots + S_T^z) + \gamma(F_1 + \delta F_1 + \cdots + \delta^{T-1}F_1] \\ &= P_0 + \theta(S_1^z + \cdots + S_T^z) + \frac{\gamma(1-\delta^T)F_1}{1-\delta}. \end{aligned} \tag{A.48}$$

Maximizing $\mathrm{E}[P_T]$ over $S_1, S_2, \ldots, S_T$ is the same as maximizing the middle term in (A.48):

$$(S_1^z + \cdots + S_T^z). \tag{A.49}$$

When $z = 1$, it does not matter how trades are allocated because (A.49) is always equal to $\bar{S}$. When $0 < z < 1$, (A.49) is a concave function with respect to $S_1, S_2, \ldots, S_T$, and is maximized when

$S_1 = S_2 = \cdots = S_T = \bar{S}/T$, which completes the proof of the optimal strategy, (A.10).

The optimal profit is simply the total value of the position subtracted by the average execution cost:

$$V^* = \mathrm{E}[P_T] \cdot \bar{S} - \mathrm{E}\left[\sum_{t=1}^{T} P_t S_t\right] = \left(\mathrm{E}[P_T] - \frac{1}{T}\sum_{t=1}^{T}\mathrm{E}[P_t]\right)\bar{S}. \tag{A.50}$$

Based on a similar derivation to (A.48), it is easy to show that

$$\mathrm{E}[P_t] = P_0 + \theta(S_1^z + \cdots + S_t^z) + \frac{\gamma(1-\delta^t)F_1}{1-\delta} = P_0 + \theta t\frac{\bar{S}^z}{T^z} + \frac{\gamma(1-\delta^t)F_1}{1-\delta}, \tag{A.51}$$

for $t = 1, 2, \ldots, T$. Substituting (A.51) into (A.50), we have

$$\begin{aligned}
V^* &= \left(\mathrm{E}[P_T] - \frac{1}{T}\sum_{t=1}^{T}\mathrm{E}[P_t]\right)\bar{S} \\
&= \left(P_0 + \theta T\frac{\bar{S}^z}{T^z} + \frac{\gamma(1-\delta^T)F_1}{1-\delta} - \frac{1}{T}\sum_{t=1}^{T}\left(P_0 + \theta t\frac{\bar{S}^z}{T^z} + \frac{\gamma(1-\delta^t)F_1}{1-\delta}\right)\right)\bar{S} \\
&= \left(\theta T\frac{\bar{S}^z}{T^z} + \frac{\gamma(1-\delta^T)F_1}{1-\delta} - \frac{1}{T}\sum_{t=1}^{T}\left(\theta t\frac{\bar{S}^z}{T^z} + \frac{\gamma(1-\delta^t)F_1}{1-\delta}\right)\right)\bar{S} \\
&= \left(\theta\frac{\bar{S}^z}{T^z}\left(T - \frac{1}{T}\sum_{t=1}^{T}t\right) + \frac{\gamma F_1}{1-\delta}\left((1-\delta^T) - \frac{1}{T}\sum_{t=1}^{T}(1-\delta^t)\right)\right)\bar{S} \\
&= \left(\theta\frac{\bar{S}^z}{T^z}\left(T - \frac{1+T}{2}\right) + \frac{\gamma F_1}{1-\delta}\left(\frac{1}{T}\sum_{t=1}^{T}\delta^t - \delta^T\right)\right)\bar{S} \\
&= \left(\theta\frac{\bar{S}^z}{T^z}\left(\frac{T-1}{2}\right) + \frac{\gamma\delta F_1}{1-\delta}\left(\frac{1-\delta^T}{T(1-\delta)} - \delta^{T-1}\right)\right)\bar{S} \\
&= \left(\frac{\theta\bar{S}^z(T-1)}{2T^z} + \frac{\gamma\delta F_1\left(1 - T\delta^{T-1} + (T-1)\delta^T\right)}{(1-\delta)^2 T}\right)\bar{S}
\end{aligned} \tag{A.52}$$

which completes the proof of (A.11).

References

Almgren, R., C. Thum, E. Hauptmann, and H. Li, 2005, Direct estimation of equity market impact, *Risk* 18, 58–62.

Bertsimas, D., and A. W. Lo, 1998, Optimal control of execution costs, *Journal of Financial Markets* 1, 1–50.

David, H. A., and H. N. Nagaraja, 2004, *Order Statistics* (John Wiley & Sons).

Fagnan, D. E., J. M. Fernandez, A. W. Lo, and R. M. Stein, 2013, Can financial engineering cure cancer?, *American Economic Review* 103, 406–11.

Irwin, D. A., and P. J. Klenow, 1996a, High-tech R&D subsidies estimating the effects of sematech, *Journal of International Economics* 40, 323–344.

Irwin, D. A., and P. J. Klenow, 1996b, Sematech: purpose and performance, *Proceedings of the National Academy of Sciences* 93, 12739–12742.

Kim, E., and A. W. Lo, 2019, Venture philanthropy: a case study of the cystic fibrosis foundation, Available at SSRN 3376673.

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

Lee, H.-M., and M. Viana, 1999, The joint covariance structure of ordered symmetrically dependent observations and their concomitants of order statistics, *Statistics & probability letters* 43, 411–414.

Lillo, F., J. D. Farmer, and R. N. Mantegna, 2003, Master curve for price-impact function, *Nature* 421, 129–130.

Link, A. N., D. J. Teece, and W. F. Finan, 1996, Estimating the benefits from collaboration: The case of sematech, *Review of Industrial Organization* 11, 737–751.

Lo, A. W., and A. C. MacKinlay, 1990, Data-snooping biases in tests of financial asset pricing models, *The Review of Financial Studies* 3, 431–467.

Pástor, L., and R. F. Stambaugh, 1999, Costs of equity capital and model mispricing, *The Journal of Finance* 54, 67–121.

Thakor, R. T., N. Anaya, Y. Zhang, C. Vilanilam, K. W. Siah, C. H. Wong, and A. W. Lo, 2017, Just how good an investment is the biopharmaceutical sector?, *Nature Biotechnology* 35, 1149–1157.

Treynor, J. L., and F. Black, 1973, How to use security analysis to improve portfolio selection, *The Journal of Business* 46, 66–86.

Yang, S.-S., 1977, General distribution theory of the concomitants of order statistics, *The Annals of Statistics* 996–1002.