Online Appendices for Cross-Sectional Variation of Intraday Liquidity, Cross-Impact, and their Effect on Portfolio Execution

Seungki Min
Graduate School of Business
Columbia University
smin20@gsb.columbia.edu

Costis Maglaras
Graduate School of Business
Columbia University
c.maglaras@gsb.columbia.edu

Ciamac C. Moallemi
Graduate School of Business
Columbia University
ciamac@gsb.columbia.edu

A. Additional Empirical Analysis

We provide the empirical analysis illustrated in §2 for the longer timespan from 2007 to 2018. The same procedure was conducted for each year and for the stocks that had been the constituents of S&P 500 throughout that year. Figure 6 shows the intraday pattern of correlation in liquidity for each year with various ways of visualization. First of all, we observe consistently across all years that the correlation increases over the course of the day. While the overall level of correlation fluctuates over years[1] (bottom left figure), we observe that the end-of-day increase had become significant (bottom right figure), which can be attributable to the increasing popularity of index-fund investing.

Figure 7 shows that the intraday pattern exists among both large-cap stocks and small-cap stocks, and we can observe that the large-cap stocks are more correlated than the small-cap stocks.

Recall that in §5.2 we have argued that the benefit from incorporate cross-asset impact into execution scheduling depends not only on the relative magnitude of index-fund liquidity provision versus single-stock liquidity provision (that is captured by the proportion of index-fund liquidity θ), but also on the intraday variation of their composition (that is captured by changes in the ratio $\frac{\alpha_t}{\beta_t}$). Table 1 shows that the maximum cost savings according to the analysis of §5 have been

[1]We observe that the correlation is relatively higher during the period 2007–2011 that coincides with the time of financial crisis. This will be consistent with a common belief that high volatility of markets is directly linked with strong correlations between stocks (Junior and Franca, 2011).

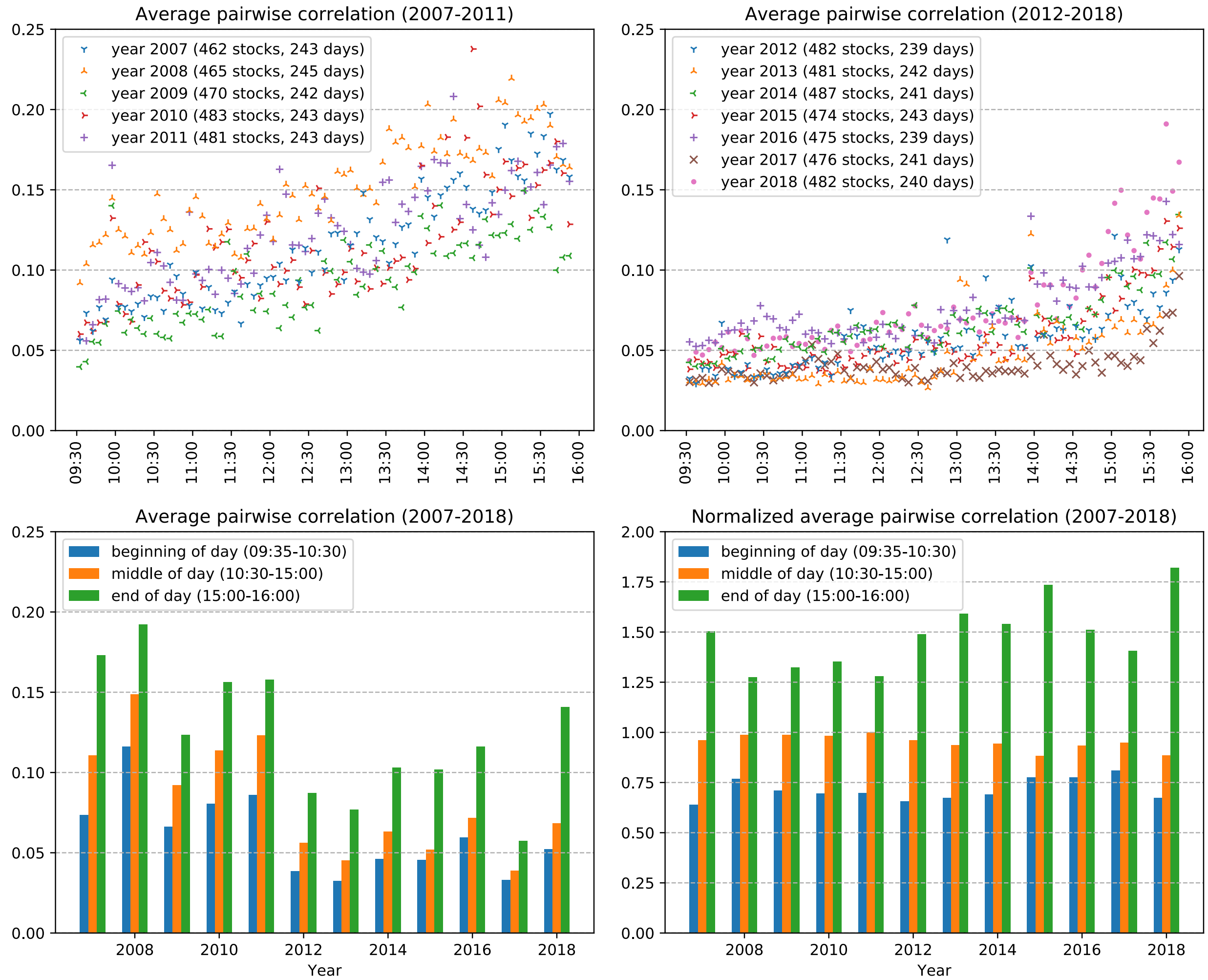

Figure 6: Intraday variations of average pairwise correlation in liquidity for years 2007–2011 (top left) and for years 2012–2018 (top right), and their alternative visualization with ones averaged for each part of day (bottom left) and ones further normalized by the average correlation level over the entire day (bottom right).

increasing in recent years: although the overall proportion of index-fund liquidity was relatively higher during the financial crisis period '07–'11, the composition of liquidity was stable throughout the day in this period, and as a result a separable VWAP execution would work fine compared to the optimized coupled execution.

B. Change of Units

In §3 and §4, the price impact was the equilibrium expected price change $\Delta\mathbf{p}$, expressed in dollars, required for the market to clear when executing a vector $\mathbf{v}$, expressed in number of shares for each security in the executed portfolio. We can restate the market impact in terms of the return $\mathbf{r} \in \mathbb{R}^N$

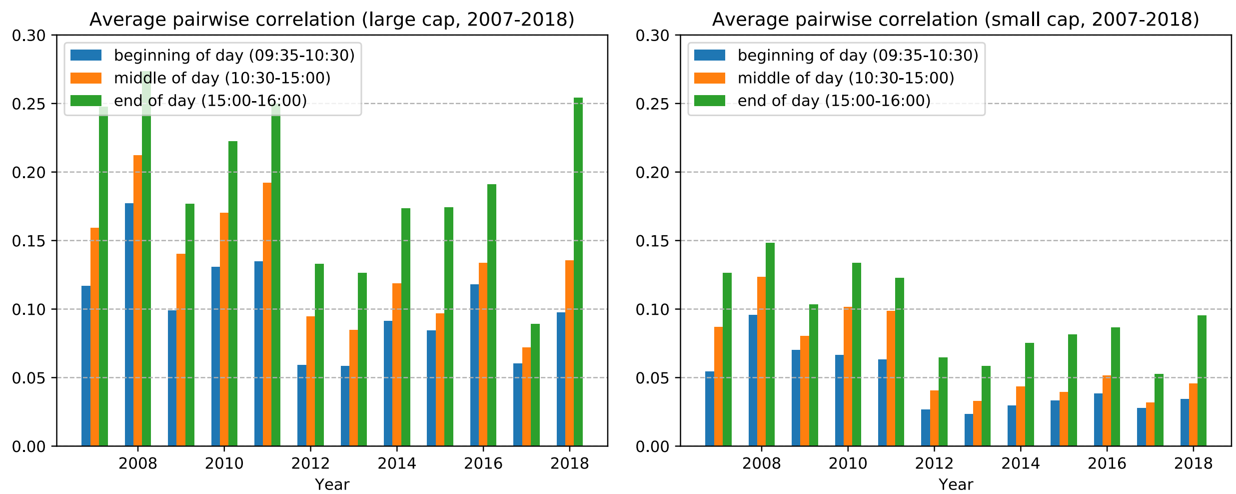

Figure 7: The average pairwise correlation for large-cap stocks (top-100 stocks in S&P 500, left) and for small-cap stocks (bottom-100 stocks in S&P 500, right).

Year	'07	'08	'09	'10	'11	'12	'13	'14	'15	'16	'17	'18
Prop. of fund liquidity (%)	26.5	29.7	24.3	26.7	27.4	20.3	19.1	21.7	20.9	22.9	17.9	23.5
Maximum cost saving (%)	8.3	3.1	7.8	7.2	6.7	11.2	16.4	8.7	12.0	6.4	7.9	14.0

Table 1: The illustrative statistics introduced in §5: the proportion of index-fund liquidity, θ, and the maximum cost saving, $\Upsilon_{\max} - 1$, estimated on years 2007–2018.

as a function of the vector of notional execution quantities $\tilde{\mathbf{v}} \in \mathbb{R}^N$. We let p denote the (arrival) equilibrium price vector $\mathbf{p} \in \mathbb{R}^N$, snapped at the beginning of the execution period, and define the diagonal matrix $\mathbf{P} \triangleq \mathrm{diag}(\mathbf{p}) \in \mathbb{R}^{N \times N}$. Then,

$$\mathbf{r} \triangleq \mathbf{P}^{-1}\Delta\mathbf{p} \quad \text{and} \quad \tilde{\mathbf{v}} \triangleq \mathbf{P}\mathbf{v}. \tag{49}$$

We redefine the liquidity variable $\psi_{\mathrm{id},i}$, $\psi_{\mathrm{f},k}$ and weight vectors $\mathbf{w}_k$ accordingly:

$$\tilde{\psi}_{\mathrm{id},i} \triangleq p_i^2 \cdot \psi_{\mathrm{id},i}, \quad \tilde{\psi}_{\mathrm{f},k} \triangleq (\mathbf{w}_k^\top \mathbf{p})^2 \cdot \psi_{\mathrm{f},k} \quad \text{and} \quad \tilde{\mathbf{w}}_k \triangleq \frac{\mathbf{P}\mathbf{w}_k}{\mathbf{p}^\top \mathbf{w}_k}. \tag{50}$$

The redefined liquidity variable $\psi_{\mathrm{id},i}$ now has the following interpretation: single-stock investors will sell (or buy) $1\% \cdot \tilde{\psi}_{\mathrm{id},it}$ dollar amount of stock i, when its price rises (or drops) by one percent. The rescaled weight vector $\tilde{\mathbf{w}}_k$ represents the normalized dollar-weighted portfolio. Putting it all together, we get

$$\begin{aligned} \mathbf{r} &= \mathbf{P}^{-1}\mathbf{G}\mathbf{v} = \mathbf{P}^{-1}\left(\boldsymbol{\Psi}_{\mathrm{id}} + \mathbf{W}\boldsymbol{\Psi}_{\mathrm{f}}\mathbf{W}^\top\right)^{-1}\mathbf{P}^{-1} \cdot \mathbf{P}\mathbf{v} \\ &= \left(\mathbf{P}\boldsymbol{\Psi}_{\mathrm{id}}\mathbf{P} + \mathbf{P}\mathbf{W}\boldsymbol{\Psi}_{\mathrm{f}}\mathbf{W}^\top\mathbf{P}\right)^{-1}\tilde{\mathbf{v}} = \left(\tilde{\boldsymbol{\Psi}}_{\mathrm{id}} + \tilde{\mathbf{W}}\tilde{\boldsymbol{\Psi}}_{\mathrm{f}}\tilde{\mathbf{W}}^\top\right)^{-1}\tilde{\mathbf{v}}. \end{aligned}$$

The resulting expected implementation shortfall cost is unchanged:

$$\bar{\mathcal{C}}(\mathbf{v}) \triangleq \frac{1}{2}\mathbf{v}^\top \Delta\mathbf{p} = \frac{1}{2}\tilde{\mathbf{v}}^\top \mathbf{r} = \frac{1}{2}\tilde{\mathbf{v}}^\top \left(\tilde{\mathbf{\Psi}}_{\mathrm{id}} + \tilde{\mathbf{W}}\tilde{\mathbf{\Psi}}_{\mathrm{f}}\tilde{\mathbf{W}}^\top\right)^{-1} \tilde{\mathbf{v}}.$$

C. Estimation of Cross-asset Market Impact

In this section, we provide a detailed description of the estimation procedure sketched in §6, and verify the procedure based on a carefully synthesized dataset.

C.1. Estimation Scheme

Required data. We assume that we have access to realized portfolio executions, their realized shortfalls, and reference information about the prevailing weight vectors of popular index funds (such as the market and sector portfolios). More specifically, we assume that the given data contains the following information. First, the portfolio transactions $\tilde{\mathbf{v}}_{dt} \in \mathbb{R}^N$ that is a portfolio vector executed during time interval t on day d, expressed in (signed) notional dollar amounts. Second, the realized implementation shortfalls (return) $\bar{\mathbf{r}}^{\mathrm{tr}}_{dt} \in \mathbb{R}^N$ incurred in the execution of portfolio $\tilde{\mathbf{v}}_{dt}$ relative to the arrival price vector at the beginning of time interval t on day d.[2] Third, some reference information that are publicly available: (i) the realized end-to-end returns $\mathbf{r}_{dt} \in \mathbb{R}^N$ during time interval t, (ii) the intraday trading volume DVol_{idt} that is the total market volume of stock i during time interval t on day d, expressed in (unsigned) notional dollar amounts, and (iii) the daily allocation of index funds $\tilde{\mathbf{W}}_d = [\tilde{\mathbf{w}}_{1d}, \ldots, \tilde{\mathbf{w}}_{Kd}] \in \mathbb{R}^{N\times K}$ where $\tilde{\mathbf{w}}_{kd}$ is the dollar-weighted vector of index fund k on day d, normalized (i.e., $\mathbf{1}^\top\tilde{\mathbf{w}}_{kd} = 1$, for all d and k). Optionally, a proxy for the amount of index-fund order flows, i.e., a quantity that reflects the trade volume generated by index-fund investors. Depending on the availability of such a proxy, we may adopt different parameterizations of the cross-impact model, (M1) or (M2), which will be introduced below.

Cross-asset impact model. The derivation in §3 predicts the following relationship between the executed quantity $\tilde{\mathbf{v}}_{dt}$ and the realized shortfall[3] $\bar{\mathbf{r}}^{\mathrm{tr}}_{dt}$: analogous to (11), we derive

$$\bar{\mathbf{r}}^{\mathrm{tr}}_{dt} = \frac{1}{2}\tilde{\mathbf{G}}_{dt}\tilde{\mathbf{v}}_{dt} + \bar{\boldsymbol{\epsilon}}^{\mathrm{tr}}_{dt}, \quad \tilde{\mathbf{G}}_{dt} = \left(\tilde{\mathbf{\Psi}}_{\mathrm{id},dt} + \tilde{\mathbf{W}}_d\tilde{\mathbf{\Psi}}_{\mathrm{f},dt}\tilde{\mathbf{W}}_d^\top\right)^{-1}, \tag{51}$$

[2] We require the return dataset $\bar{\mathbf{r}}^{\mathrm{tr}}_{dt}$ to have no missing entries. For an entry (i, d, t) such that no execution was made at all (i.e., $\tilde{v}_{idt} = 0$), we recommend to set $\bar{r}^{\mathrm{tr}}_{idt}$ to be the return measured with the market volume-weighted-average-price relative to the arrival price at the beginning of interval, or simply a half of the end-to-end market return $\frac{1}{2}r_{idt}$.

[3] In this section, we are using the realized shortfall (return) $\bar{\mathbf{r}}^{\mathrm{tr}}_{dt}$ instead of the absolute price change $\Delta\mathbf{p}_{dt}$, and notional traded vectors $\tilde{\mathbf{v}}_{dt}$ instead of number of shares $\mathbf{v}_{dt}$. Similarly, we use dollar-weighted vectors $\tilde{\mathbf{w}}_k$ instead of share-weighted vectors $\mathbf{w}_k$. With the rescaled liquidity parameters $\tilde{\mathbf{\Psi}}_{\mathrm{id}}$ and $\tilde{\mathbf{\Psi}}_{\mathrm{f}}$, the structure of the price-impact model remains the same. See Appendix B.

where the rescaled liquidity matrices are given by $\tilde{\boldsymbol{\Psi}}_{\mathrm{id},dt} = \mathrm{diag}_{i=1}^{N}(\tilde{\psi}_{\mathrm{id},idt}) \in \mathbb{R}^{N\times N}$ and $\tilde{\boldsymbol{\Psi}}_{\mathrm{f},dt} = \mathrm{diag}_{k=1}^{K}(\tilde{\psi}_{\mathrm{f},kdt}) \in \mathbb{R}^{K\times K}$, and the noise term $\bar{\boldsymbol{\epsilon}}_{dt}^{\mathrm{tr}} \in \mathbb{R}^N$ describes the random fluctuation of price. Next we introduce a further parameterization of $\tilde{\boldsymbol{\Psi}}_{\mathrm{id},dt}$ and $\tilde{\boldsymbol{\Psi}}_{\mathrm{f},dt}$, in which we reduce the number of free parameters for the idiosyncratic components (as is typically done), and further simplify how we capture the non-stationary behavior of the various terms so as to be able to rely on market observable quantities as proxies.

A parameterization with "idiosyncratic" and "factor" trading volume. As discussed in §3, the liquidity variable $\tilde{\psi}_{\mathrm{id},idt}$ (resp., $\tilde{\psi}_{\mathrm{f},kdt}$) represents the notional amount of stock i (resp., index fund k) that will be supplied by single-stock investors (resp., index-fund investors) in response to a movement in the price of the stock (resp., index). We interpret that the variable $\tilde{\psi}$ captures (i) the number of investors, or participation intensity, present in each period, and (ii) the sensitivity of these investors to price movements. The first factor roughly scales in proportion to trading volume, while the second factor varies in a way that depends on the volatility of the underlying security or index, and, specifically, it is plausible to imagine that it scales in a way that it is inversely proportional to the volatility itself, i.e., $\tilde{\psi} \propto \frac{\mathrm{DVol}}{\sigma}$. Based on this interpretation, we consider the parameterizations of $\tilde{\psi}_{\mathrm{id},idt}$ and $\tilde{\psi}_{\mathrm{f},kdt}$ with the following reduced form:

$$\tilde{\psi}_{\mathrm{id},idt} = \gamma_{\mathrm{id}} \times \frac{\widehat{\mathrm{DVol}}_{\mathrm{id},idt}}{\widehat{\sigma}_{\mathrm{id},idt}}, \quad \tilde{\psi}_{\mathrm{f},kdt} = \gamma_{\mathrm{f},k} \times \frac{\widehat{\mathrm{DVol}}_{\mathrm{f},kdt}}{\widehat{\sigma}_{\mathrm{f},kdt}}, \tag{M1}$$

where (i) $\widehat{\mathrm{DVol}}_{\mathrm{id},idt}$ and $\widehat{\sigma}_{\mathrm{id},idt}$ denote (forecasted) "idiosyncratic" trading volume and volatility of stock i that describe the trading activity of single-stock investors, (ii) $\widehat{\mathrm{DVol}}_{\mathrm{f},kdt}$ and $\widehat{\sigma}_{\mathrm{f},kdt}$ denote (forecasted) "factor" trading volume and volatility of index fund k that describe the trading activity of index-fund investors, and (iii) $\gamma_{\mathrm{id}} \in \mathbb{R}$ and $\boldsymbol{\gamma}_{\mathrm{f}} \triangleq (\gamma_{\mathrm{f},1}, \ldots, \gamma_{\mathrm{f},K})^\top \in \mathbb{R}^K$ are unknown time-invariant leading coefficients. We have selected a simple parameterization where all single-stock terms $\tilde{\psi}_{\mathrm{id},idt}$ share the same coefficient γ_{id} that is believed to reflect some invariant characteristic of all single-stock investors.

We do not further formulate "idiosyncratic" and "factor" trading volumes in this paper: they will be latent variables, i.e., they are not immediately quantifiable from the market data, since the actual trading volume that we observe from the market is the mixture of these two components. Someone can estimate the intraday pattern of the proportion of factor trading volume explicitly as suggested in §5.1, or can use some additional market information such as trading volume grouped by investor type if available. Given such proxies, $K+1$ unknown parameters (γ_{id} and $\boldsymbol{\gamma}_{\mathrm{f}}$) can be estimated via a procedure described later.

A parameterization with observables. As an alternative of the parameterization (M1), we propose a more specific parameterization that relies on directly observable quantities:

$$\tilde{\psi}_{\mathrm{id},idt} = \nu_{\mathrm{id},t} \times \frac{\mathrm{MADVol}_{idt}}{\bar{\sigma}_{idt}}, \quad \tilde{\psi}_{\mathrm{f},kdt} = \nu_{\mathrm{f},kt} \times \frac{\sum_{i\in\mathcal{S}_k}\mathrm{MADVol}_{idt}}{\bar{\sigma}_{\mathrm{f},kdt}}. \tag{M2}$$

The coefficients $\nu_{\mathrm{id},t}$ and $\boldsymbol{\nu}_{\mathrm{f},t}$ are the unknowns here, and the others variables are the moving-average measures defined as

$$\mathrm{MADVol}_{idt} \triangleq \frac{1}{\tau}\sum_{s=0}^{\tau-1}\mathrm{DVol}_{i,d-s,t}, \quad \bar{\sigma}_{idt} \triangleq \sqrt{\frac{1}{\tau}\sum_{s=0}^{\tau-1} r_{i,d-s,t}^2}, \quad \bar{\sigma}_{\mathrm{f},kdt} \triangleq \sqrt{\frac{1}{\tau}\sum_{s=0}^{\tau-1}\left(\tilde{\mathbf{w}}_{k,d-s}^{\top}\mathbf{r}_{d-s,t}\right)^2}, \tag{52}$$

where τ is the length of sliding window and $\mathcal{S}_k$ is the set of stocks that belong to the index fund k. One can adopt different averaging scheme as long as it provides reasonable forecasts for trading volume and volatility.

Compared to the previous parameterization (M1), the leading coefficients in (M2), $\nu_{\mathrm{id},t}$ and $\boldsymbol{\nu}_{\mathrm{f},t}$, have the subscript t (as opposed to γ_{id} and $\boldsymbol{\gamma}_{\mathrm{f}}$) so as to reflect the intraday variation in composition of the two types of liquidity provision (i.e., $\tilde{\psi}_{\mathrm{id},idt}$ vs. $\tilde{\psi}_{\mathrm{f},kdt}$) correctly. Such a variation is not well captured in the moving-averaged measures introduced above (i.e., MADVol_{idt} vs. $\sum_{i\in\mathcal{S}_k}\mathrm{MADVol}_{idt}$), since the observable trading volume DVol_{idt} is a simple reflection of sum of two types of liquidity.[4] By allowing the unknown coefficients dependent on the time of the day, we let the estimation procedure to find the right values of $\nu_{\mathrm{id},t}$ and $\boldsymbol{\nu}_{\mathrm{f},t}$ that fairly describe the expected intraday profile of liquidity composition.

There are $T\times(K+1)$ values to estimate, too many considering the noise level in the dataset. We further reduce the number of unknowns by imposing a simple intraday variation pattern: we divide a day into three segments and assume that the coefficients are constant within each segment. More specifically, let $\mathcal{T}_{\mathrm{beg}}$, $\mathcal{T}_{\mathrm{end}}$, and $\mathcal{T}_{\mathrm{mid}}$ be the first one hour (09:30–10:30), the last one hour (15:00–16:00), and the remaining trading session (10:30–15:00), respectively, and assume that

$$\nu_{\mathrm{id},t} = \begin{cases} \nu_{\mathrm{id}}^{\mathrm{beg}} & \text{if } t\in\mathcal{T}_{\mathrm{beg}} \\ \nu_{\mathrm{id}}^{\mathrm{mid}} & \text{if } t\in\mathcal{T}_{\mathrm{mid}} \\ \nu_{\mathrm{id}}^{\mathrm{end}} & \text{if } t\in\mathcal{T}_{\mathrm{end}} \end{cases}, \quad \nu_{\mathrm{f},kt} = \begin{cases} \nu_{\mathrm{f},k}^{\mathrm{beg}} & \text{if } t\in\mathcal{T}_{\mathrm{beg}} \\ \nu_{\mathrm{f},k}^{\mathrm{mid}} & \text{if } t\in\mathcal{T}_{\mathrm{mid}} \\ \nu_{\mathrm{f},k}^{\mathrm{end}} & \text{if } t\in\mathcal{T}_{\mathrm{end}} \end{cases}. \tag{M2-seg}$$

With this segmentation, we have $3\times(K+1)$ unknowns in total, and the intraday variation in

[4] Suppose that the intensity of liquidity provision by index-fund investors stays constant over time, i.e., $\tilde{\psi}_{\mathrm{f},kdt}$ does not vary over the course of the day but does $\tilde{\psi}_{\mathrm{id},idt}$ only. The market-wide volume $\sum_{i\in\mathcal{S}_k}\mathrm{MADVol}_{idt}$ will still fluctuate according to the variation of single-stock investors' liquidity provision, and therefore, the time-variation of $\sum_{i\in\mathcal{S}_k}\mathrm{MADVol}_{idt}$ does not correctly reflect the time-variation of index-fund investors' liquidity provision.

liquidity composition can be represented with the change in their relative magnitude across the segments, i.e., $\nu_{\text{id}}^{\text{beg}}$ vs. $\boldsymbol{\nu}_{\text{f}}^{\text{beg}}$, $\nu_{\text{id}}^{\text{mid}}$ vs. $\boldsymbol{\nu}_{\text{f}}^{\text{mid}}$, and $\nu_{\text{id}}^{\text{end}}$ vs. $\boldsymbol{\nu}_{\text{f}}^{\text{end}}$.

Estimation procedure. We illustrate a simple procedure that estimates the unknown coefficients in the parameterization (M2) for intraday segment $\mathcal{T}_{\text{end}}$. The same procedure can apply for the other intraday segments as well as the parameterization (M1).

We aim to find the values of $\nu_{\text{id}}^{\text{end}} \in \mathbb{R}$ and $\boldsymbol{\nu}_{\text{f}}^{\text{end}} \in \mathbb{R}^K$ such that their corresponding cross-impact model fits the actual price changes realized during the time periods $t \in \mathcal{T}_{\text{end}}$. Let us denote the cross-impact matrix parameterized with ν_{id} and $\boldsymbol{\nu}_{\text{f}}$ by $\tilde{\mathbf{G}}_{dt}(\nu_{\text{id}}, \boldsymbol{\nu}_{\text{f}})$: More specifically,

$$\tilde{\mathbf{G}}_{dt}(\nu_{\text{id}}, \boldsymbol{\nu}_{\text{f}}) \triangleq \left(\text{diag}_{i=1}^{N} \left(\nu_{\text{id}} \cdot \frac{\text{MADVol}_{idt}}{\bar{\sigma}_{idt}} \right) + \tilde{\mathbf{W}}_d \text{diag}_{k=1}^{K} \left(\nu_{\text{f},k} \cdot \frac{\sum_{i \in \mathcal{S}_k} \text{MADVol}_{idt}}{\bar{\sigma}_{\text{f},kdt}} \right) \tilde{\mathbf{W}}_d^\top \right)^{-1}.$$

We first introduce the empirical loss $\mathcal{L}_{\text{id}}^{\text{end}}$ with respect to the realized single-stock shortfalls $\bar{\mathbf{r}}_{dt}^{\text{tr}}$: as we expect that $\bar{\mathbf{r}}_{dt}^{\text{tr}} \approx \frac{1}{2} \tilde{\mathbf{G}}_{dt}(\nu_{\text{id}}, \boldsymbol{\nu}_{\text{f}}) \tilde{\mathbf{v}}_{dt}$,

$$\mathcal{L}_{\text{id}}^{\text{end}}(\nu_{\text{id}}, \boldsymbol{\nu}_{\text{f}}) \triangleq \frac{1}{N} \sum_{d=1}^{D} \sum_{t \in \mathcal{T}_{\text{end}}} \left(\bar{\mathbf{r}}_{dt}^{\text{tr}} - \frac{1}{2} \tilde{\mathbf{G}}_{dt}(\nu_{\text{id}}, \boldsymbol{\nu}_{\text{f}}) \tilde{\mathbf{v}}_{dt} \right)^\top \bar{\boldsymbol{\Sigma}}_{dt}^{-1} \left(\bar{\mathbf{r}}_{dt}^{\text{tr}} - \frac{1}{2} \tilde{\mathbf{G}}_{dt}(\nu_{\text{id}}, \boldsymbol{\nu}_{\text{f}}) \tilde{\mathbf{v}}_{dt} \right),$$

where $\bar{\boldsymbol{\Sigma}}_{dt} \triangleq \text{diag}_{i=1}^{N}(\bar{\sigma}_{idt}^2)$ is the empirical diagonal covariance matrix. Analogously, the empirical loss $\mathcal{L}_{\text{f}}^{\text{end}}$ with respect to the realized index-fund shortfalls $\tilde{\mathbf{W}}_d^\top \bar{\mathbf{r}}_{dt}^{\text{tr}}$ can be defined as follows:

$$\mathcal{L}_{\text{f}}^{\text{end}}(\nu_{\text{id}}, \boldsymbol{\nu}_{\text{f}}) \triangleq \frac{1}{K} \sum_{d=1}^{D} \sum_{t \in \mathcal{T}_{\text{end}}} \left(\tilde{\mathbf{W}}_d^\top \bar{\mathbf{r}}_{dt}^{\text{tr}} - \frac{1}{2} \tilde{\mathbf{W}}_d^\top \tilde{\mathbf{G}}_{dt}(\nu_{\text{id}}, \boldsymbol{\nu}_{\text{f}}) \tilde{\mathbf{v}}_{dt} \right)^\top \bar{\boldsymbol{\Sigma}}_{\text{f},dt}^{-1} \left(\tilde{\mathbf{W}}_d^\top \bar{\mathbf{r}}_{dt}^{\text{tr}} - \frac{1}{2} \tilde{\mathbf{W}}_d^\top \tilde{\mathbf{G}}_{dt}(\nu_{\text{id}}, \boldsymbol{\nu}_{\text{f}}) \tilde{\mathbf{v}}_{dt} \right),$$

where $\bar{\boldsymbol{\Sigma}}_{\text{f},dt} \triangleq \text{diag}_{k=1}^{K}(\bar{\sigma}_{kdt}^2)$ is the empirical diagonal covariance matrix of index-fund returns.

Observe that minimizing $\mathcal{L}_{\text{id}}^{\text{end}}$ or $\mathcal{L}_{\text{f}}^{\text{end}}$ is identical to performing a least squares estimation with the heteroscedastic residual terms. In particular when the index-fund investors' contribution is absent (i.e., when we are restricted to have $\boldsymbol{\nu}_{\text{f}} = \mathbf{0}$), minimizing $\mathcal{L}_{\text{id}}^{\text{end}}$ is equivalent to the estimation procedure proposed in Almgren et al. (2005). Similarly, when the single-stock investors' contribution is absent (i.e., when we are restricted to have $\nu_{\text{id}} = 0$) and the index funds are orthogonal, minimizing $\mathcal{L}_{\text{f}}^{\text{end}}$ is equivalent to fitting a separable linear model under which each index fund is treated in isolation. In other words, the loss $\mathcal{L}_{\text{id}}^{\text{end}}$ focuses more on the diagonal entries of cross-impact matrix $\tilde{\mathbf{G}}_{dt}$ and the loss $\mathcal{L}_{\text{f}}$ rather focuses on the non-diagonal entries of $\tilde{\mathbf{G}}_{dt}$.

Based on those loss measures, we suggest a four-step procedure that yields the estimates $\widehat{\nu}_{\text{id}}^{\text{end}}$ and $\widehat{\boldsymbol{\nu}}_{\text{f}}^{\text{end}}$:

1. (Initial guess) Find a single scalar value $\widehat{\nu}$ via an ordinary least-squares regression based on the following linear model:

$$\bar{r}^{\text{tr}}_{idt} = \frac{1}{2} \times \nu^{-1} \times \frac{\bar{\sigma}_{idt}}{\text{MADVol}_{idt}} \times \tilde{v}_{idt} + e_{idt},$$

where e_{idt}'s are i.i.d. Initialize the estimates with $\widehat{\nu}$: i.e., $\widehat{\nu}^{\text{end}}_{\text{id}} \leftarrow \widehat{\nu}$ and $\widehat{\nu}^{\text{end}}_{\text{f},k} \leftarrow \widehat{\nu}$ for all $k = 1, \ldots, K$.

2. (Estimation of diagonal entries) Fix $\widehat{\boldsymbol{\nu}}^{\text{end}}_{\text{f}}$ and find $\widehat{\nu}^{\text{end}}_{\text{id}}$ that best explains the realized single-stock shortfalls by minimizing loss $\mathcal{L}^{\text{end}}_{\text{id}}$:

$$\widehat{\nu}^{\text{end}}_{\text{id}} \leftarrow \underset{\nu_{\text{id}} \in \mathbb{R}_+}{\text{argmin}}\, \mathcal{L}^{\text{end}}_{\text{id}}(\nu_{\text{id}}, \widehat{\boldsymbol{\nu}}^{\text{end}}_{\text{f}}).$$

3. (Estimation of non-diagonal entries) Fix $\widehat{\nu}^{\text{end}}_{\text{id}}$ and find $\widehat{\boldsymbol{\nu}}^{\text{end}}_{\text{f}}$ that best explains the realized index-fund shortfalls by minimizing loss $\mathcal{L}^{\text{end}}_{\text{f}}$:

$$\widehat{\boldsymbol{\nu}}^{\text{end}}_{\text{f}} \leftarrow \underset{\boldsymbol{\nu}_{\text{f}} \in \mathbb{R}^K_+}{\text{argmin}}\, \mathcal{L}^{\text{end}}_{\text{f}}(\widehat{\nu}^{\text{end}}_{\text{id}}, \boldsymbol{\nu}_{\text{f}}).$$

4. (Fine tuning) Finally adjust the estimates by minimizing two loss measures simultaneously:

$$(\widehat{\nu}^{\text{end}}_{\text{id}}, \widehat{\boldsymbol{\nu}}^{\text{end}}_{\text{f}}) \leftarrow \underset{(\nu_{\text{id}}, \boldsymbol{\nu}_{\text{f}}) \in \mathbb{R}_+ \times \mathbb{R}^K_+}{\text{argmin}} \left\{ \mathcal{L}^{\text{end}}_{\text{id}}(\nu_{\text{id}}, \boldsymbol{\nu}_{\text{f}}) + \mathcal{L}^{\text{end}}_{\text{f}}(\nu_{\text{id}}, \boldsymbol{\nu}_{\text{f}}) \right\}$$

Step 1 replicates the procedure to estimate a separable (idiosyncratic) impact model that is commonly adopted in the literature, and the estimate $\widehat{\nu}$ found in step 1 is utilized as a baseline value for $\widehat{\nu}^{\text{end}}_{\text{id}}$ and $\widehat{\boldsymbol{\nu}}^{\text{end}}_{\text{f}}$ in the next steps. In steps 2 and 3, it performs further estimation of $\widehat{\nu}^{\text{end}}_{\text{id}}$ and $\widehat{\boldsymbol{\nu}}^{\text{end}}_{\text{f}}$ by minimizing the losses $\mathcal{L}^{\text{end}}_{\text{id}}$ and $\mathcal{L}^{\text{end}}_{\text{f}}$ individually, and in step 4 it performs a fine tuning by minimizing $\mathcal{L}^{\text{end}}_{\text{id}} + \mathcal{L}^{\text{end}}_{\text{f}}$ together. In the implementation, we suggest to use a simple gradient descent method in each step: the loss minimizer needs not to be the global optimum since the results from the previous steps will provide reasonable initial solutions to the next steps.

This four-step procedure aims to estimate the coefficients $\nu^{\text{end}}_{\text{id}}$ and $\boldsymbol{\nu}^{\text{end}}_{\text{f}}$ in a robust and efficient way. By sequentially improving the estimates starting from the parameter value estimated from a simple separable impact model, it prevents the final outcomes from taking extreme values and accelerates the optimization procedure. One may performs step 4 only, but we anticipate that the outcome will be very sensitive to initialization values because the objective is non-convex and possibly multimodal. The matrix $\tilde{\mathbf{G}}_{dt}(\nu_{\mathrm{id}}, \boldsymbol{\nu}_{\mathrm{f}})$ can be computed efficiently in practice by using the Woodbury matrix identity.

Comparison with Schneider and Lillo (2019). Schneider and Lillo (2019) perform a direct non-parametric estimation of cross-impact among Italian and European bonds ($N = 33$) using high-frequency market data in an effort to validate their theoretical findings on no-arbitrage conditions. Compared to their estimation procedure, our estimation heavily relies on the model: we postulate a parsimonious parametric representation of cross-impact and exploit its structure to alleviate the difficulty of direct estimation. Since our model is based on a stylized characterization of index-fund investors participating the stock markets, our suggested estimation scheme may not be appropriate to estimate cross-impact among fixed-income securities that may share a different kind of commonality following from risk and term structure. On the other hand, their non-parametric approach may not be appropriate to handle non-stationary cross-impact as opposed to our parametric approach that uses trading volume as a proxy for time-varying liquidity. Although it would be an interesting research topic to adopt direct estimation methods and compare the results, we believe that it would be beyond the scope of this paper.

C.2. Illustration with Synthetic Data

We illustrate and verify the suggested estimation procedure by using a (partly) synthetic dataset. We construct test portfolios, one for each day d and period t. We will then simulate the execution of these portfolios by adding some market impact on top of the actual market return in (d, t). We pick randomly a set of market impact coefficients for the model in (M1), and simulate execution costs for the test portfolios by adding the expected impact cost contribution to the realized market return. We then forget the impact cost coefficients and the detailed specification of (M1), and given the set of test portfolios and realized execution costs we estimate an impact model given the (observable) parameterization (M2). Even though we use slightly different market impact models for the dataset generation and for the parameter estimation, we will illustrate that it can still predict the transaction costs well enough in spite of the model misspecification.

Original dataset. We consider three sectors ($K = 3$; energy, finance, and technology sectors) and S&P 500 stocks that belong to these sectors ($N = 153$) throughout the year 2018. As in §2, we consider five-minute intervals ($T = 78$) per each day, and exclude half-trading days and the days on which FED announcements were made. In the calculation of moving-averaged measures (52), we use the time window of $\tau = 60$ days, and therefore the estimation is performed after excluding the first 60 days ($D = 240 - 60 = 180$).

Ground truth market impact model. We assume that the ground truth model is given by the parameterization (M1) where the idiosyncratic volume forecast $\widehat{\text{DVol}}_{\text{id},idt}$ and the factor volume forecast $\widehat{\text{DVol}}_{\text{f},kdt}$ are assumed to have the following form:

$$\widehat{\text{DVol}}_{\text{f},kdt} = \theta_{kt} \times \sum_{i \in \mathcal{S}_k} \text{MADVol}_{idt}, \quad \widehat{\text{DVol}}_{\text{id},idt} = \text{MADVol}_{idt} - w_{kdi} \times \widehat{\text{DVol}}_{\text{f},kdt}. \tag{53}$$

The variable θ_{kt} represents the proportion of the factor trading volume out of total market trading volume across all stocks in sector k on the intraday time interval t.[5] This formulation follows from the assumption that individual stock's trading volume is decomposed into an idiosyncratic component and a factor component, i.e., $\text{MADVol}_{idt} = \widehat{\text{DVol}}_{\text{id},idt} + w_{kdi} \times \widehat{\text{DVol}}_{\text{f},kdt}$, where the sector-wide contribution of factor volume, $\frac{\widehat{\text{DVol}}_{\text{f},kdt}}{\sum_{i \in \mathcal{S}_k} \text{MADVol}_{idt}}$, is assumed to be constant across days. Reflecting the empirical observations in §2, we make up the values of θ_{kt}'s as plotted in Figure 8.

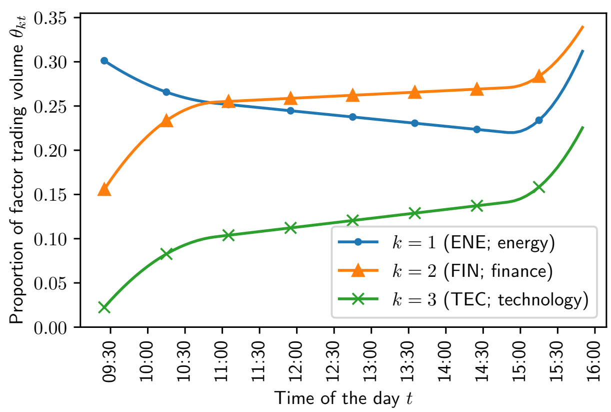

Figure 8: Hypothetical intraday profiles of the proportion of factor trading volume, $(\theta_{kt})_{t=1}^{T}$, that are plugged in (53) for synthetic dataset generation. For example, the second profile curve implies that the index-fund investors in finance sector ($k = 2$) account for 15% of the total sector-wide traded volume at the beginning of the day, and 35% of it at the end of the day.

The values of unknown coefficients γ's are chosen as $\gamma_{\text{id}} = 0.04$ and $\boldsymbol{\gamma}_{\text{f}} = (0.30, 0.10, 0.20)^\top$. To gain some intuition of the magnitude of these parameters, when $\gamma = 0.04$ and $\sigma = 0.1\% \approx \frac{1\%}{\sqrt{78}}$ (five-minute volatility), the given parameterization predicts that the expected cost of executing a trade with 2% participation rate will be $\frac{1}{2} \cdot \frac{\sigma}{\gamma} \cdot 2\% \approx 2.5$ basis points.

Given the hypothetical values of θ's and γ's, we assume that the true coefficient matrix of market

[5] Compared to the parameterization introduced in §5, the variable θ_{kt} would correspond to the proportion of index-fund order flows, $\frac{\beta_t \cdot \theta}{\alpha_t \cdot (1-\theta) + \beta_t \cdot \theta}$, which is plotted in Figure 3.

impact is given by

$$\tilde{\mathbf{G}}_{dt}^{\text{true}} = \left(\text{diag}_{i=1}^{N} \left(\gamma_{\text{id}} \cdot \frac{\widehat{\text{DVol}}_{\text{id},idt}}{\bar{\sigma}_{idt}} \right) + \tilde{\mathbf{W}}_d \text{diag}_{k=1}^{K} \left(\gamma_{\text{f},k} \cdot \frac{\widehat{\text{DVol}}_{\text{f},kdt}}{\bar{\sigma}_{\text{f},kdt}} \right) \tilde{\mathbf{W}}_d^{\top} \right)^{-1}. \tag{54}$$

In what follows, we show that our proposed estimation procedure finds some approximation of $\tilde{\mathbf{G}}_{dt}^{\text{true}}$ with a different parameterization rather than directly estimating the values of θ's and γ's.

Hypothetical portfolio transactions. We imagine a situation that investing decisions are made on a daily basis and the associated portfolio transactions are being executed over the course of the day. More specifically, the hypothetical portfolio transactions $\tilde{\mathbf{v}}_{dt}$'s are generated according to the following procedure: on each day $d = 1, \ldots, D$, independently,

1. we randomly select the single stocks and the sectors to trade: A single stock is selected with probability 5%, and a sector is selected with probability 25%.
2. For each selected single stock (or a selected sector), the trading direction (i.e., buy or sell) is determined randomly, and the participation rate is drawn from the log-normal distribution with mean 1.5% and standard deviation 0.5% for single stocks, and with mean 0.5% and standard deviation 0.1% for sectors.
3. Given the trading direction and the participation rate, we imagine a VWAP trading schedule over the course of the day: The notional amount of transactions at time t on a selected asset (a stock i or a sector k) is given by

$$\begin{aligned} q_{\text{id},idt} &= (\text{trading direction})_{id} \times (\text{participation rate})_{id} \times \text{MADVol}_{idt}, \\ q_{\text{f},kdt} &= (\text{trading direction})_{kd} \times (\text{participation rate})_{kd} \times \sum_{i \in \mathcal{S}_k} \text{MADVol}_{idt}. \end{aligned}$$

   Accordingly, the portfolio transactions on day d are simply given by

$$\tilde{\mathbf{v}}_{dt} = \mathbf{q}_{\text{id},dt} + \tilde{\mathbf{W}}_d \mathbf{q}_{\text{f},dt}, \quad \forall t = 1, \ldots, T,$$

   where $\mathbf{q}_{\text{id},dt} \triangleq (q_{\text{id},1dt}, \ldots, q_{\text{id},Ndt})^{\top} \in \mathbb{R}^N$ and $\mathbf{q}_{\text{f},dt} \triangleq (q_{\text{f},1dt}, \ldots, q_{\text{f},Kdt})^{\top} \in \mathbb{R}^K$.

This procedure is selected intentionally to match up to calibrating the model on a set of full day VWAP-like executions.

Hypothetical dataset. Given the ground truth impact model $\tilde{\mathbf{G}}_{dt}^{\text{true}}$ and the simulated portfolio transactions $\tilde{\mathbf{v}}_{dt}$, we make perturbation on the original dataset: The realized implementation shortfalls $\bar{\mathbf{r}}_{dt}^{\text{tr}}$, the realized end-to-end returns $\mathbf{r}_{dt}$ and the market trading volume DVol_{idt} are overwritten as

$$\bar{\mathbf{r}}^{\mathrm{tr}}_{dt} \leftarrow \bar{\mathbf{r}}^{\mathrm{tr}}_{dt} + \frac{1}{2}\tilde{\mathbf{G}}^{\mathrm{true}}_{dt}\tilde{\mathbf{v}}_{dt}, \quad \mathbf{r}_{dt} \leftarrow \mathbf{r}_{dt} + \tilde{\mathbf{G}}^{\mathrm{true}}_{dt}\tilde{\mathbf{v}}_{dt}, \quad \mathrm{DVol}_{idt} \leftarrow \mathrm{DVol}_{idt} + |\tilde{v}_{idt}|, \tag{55}$$

and we obtain a hypothetical dataset in which the effects of the transactions $\tilde{\mathbf{v}}_{dt}$ are reflected. In this hypothetical dataset, 59% of our trades are due to index-fund investing, and we pay 2.4 basis points for the transaction cost in total.

Estimation result. We perform the estimation based on the parameterization (M2) with the segmentation scheme (M2-seg). We estimate $3\times(K+1)$ unknown coefficients $(\widehat{\nu}^{\mathrm{beg}}_{\mathrm{id}}, \widehat{\boldsymbol{\nu}}^{\mathrm{beg}}_{\mathrm{f}})$, $(\widehat{\nu}^{\mathrm{mid}}_{\mathrm{id}}, \widehat{\boldsymbol{\nu}}^{\mathrm{mid}}_{\mathrm{f}})$, and $(\widehat{\nu}^{\mathrm{end}}_{\mathrm{id}}, \widehat{\boldsymbol{\nu}}^{\mathrm{end}}_{\mathrm{f}})$ by applying the four-step procedure described in §C.1 for each of intraday segments $\mathcal{T}_{\mathrm{beg}}$, $\mathcal{T}_{\mathrm{mid}}$, and $\mathcal{T}_{\mathrm{end}}$ separately.

Since the estimation model is different from the model used for data generation, there are no true parameter values that exactly correspond to the estimated parameters. Instead, we introduce the "effective" true coefficients for (M2) that are expressed in terms of θ and γ used in (M1):

$$\bar{\nu}_{\mathrm{id},t} \triangleq \gamma_{\mathrm{id}} \times \left(1 - \frac{1}{K}\sum_{k=1}^{K}\theta_{kt}\right), \quad \bar{\nu}_{\mathrm{f},kt} \triangleq \gamma_{\mathrm{f},k} \times \theta_{kt}.$$

Figure 9 (left) shows the comparison between the estimated coefficients $(\widehat{\nu}_{\mathrm{id},t}, \widehat{\boldsymbol{\nu}}_{\mathrm{f},t})_{t=1}^{T}$ and their effective true values $(\bar{\nu}_{\mathrm{id},t}, \bar{\boldsymbol{\nu}}_{\mathrm{f},t})_{t=1}^{T}$. We observe that the estimated model approximates the intraday variation of the ground truth market impact with piecewise constant profiles.

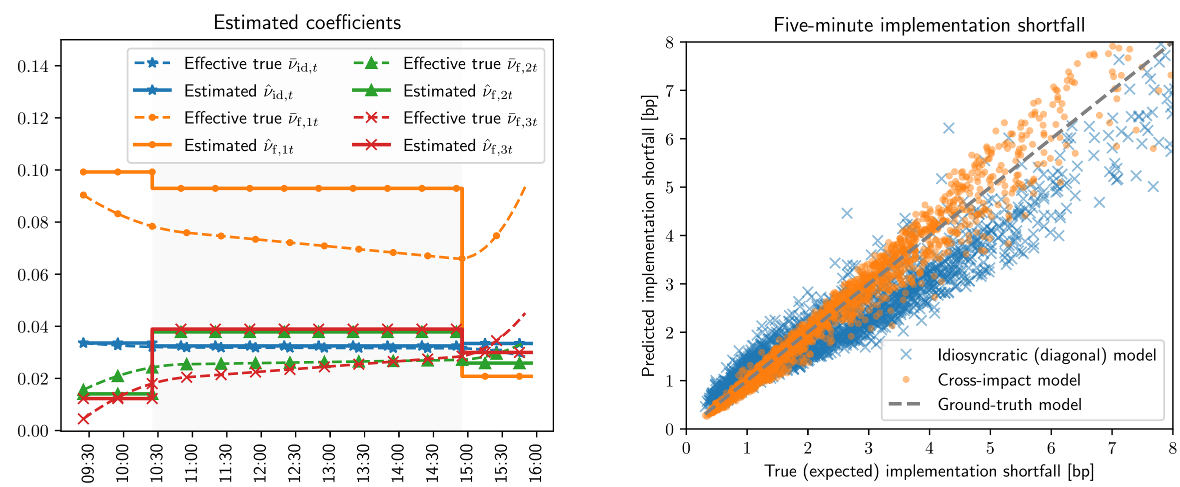

Figure 9: Estimated coefficients $\widehat{\boldsymbol{\nu}}$ vs. effective true coefficients $\bar{\boldsymbol{\nu}}$ (left). The estimation results approximate the intraday variation of the cross-impact matrix. The estimated five-minute implementation shortfall $\hat{y}^{\mathrm{cross}}_{dt}$ and $\hat{y}^{\mathrm{diag}}_{dt}$ vs. the true expected five-minute implementation shortfall $\hat{y}^{\mathrm{true}}_{dt}$ (right, the data points are randomly selected for visualization).

We further investigate the performance of estimated model relative to the idiosyncratic (diagonal) market impact model[6] that may be adopted by someone who ignores the cross-asset impact. More specifically, we compute the realized five-minute implementation shortfall y_{dt}^{real} (expressed in percentage) incurred by the hypothetical portfolio transaction $\tilde{\mathbf{v}}_{dt}$ and the model predictions:

$$y_{dt}^{\text{real}} \triangleq \frac{\bar{\mathbf{r}}_{dt}^{\top}\tilde{\mathbf{v}}_{dt}}{\|\tilde{\mathbf{v}}_{dt}\|_1},\ \hat{y}_{dt}^{\text{true}} \triangleq \frac{\frac{1}{2}\tilde{\mathbf{v}}_{dt}^{\top}\tilde{\mathbf{G}}_{dt}^{\text{true}}\tilde{\mathbf{v}}_{dt}}{\|\tilde{\mathbf{v}}_{dt}\|_1},\ \hat{y}_{dt}^{\text{cross}} \triangleq \frac{\frac{1}{2}\tilde{\mathbf{v}}_{dt}^{\top}\tilde{\mathbf{G}}_{dt}(\widehat{\nu}_{\text{id},t},\widehat{\boldsymbol{\nu}}_{\text{f},t})\tilde{\mathbf{v}}_{dt}}{\|\tilde{\mathbf{v}}_{dt}\|_1},\ \hat{y}_{dt}^{\text{diag}} \triangleq \frac{\frac{1}{2}\tilde{\mathbf{v}}_{dt}^{\top}\tilde{\mathbf{G}}_{dt}(\widehat{\nu}_{t}^{\text{diag}},\mathbf{0})\tilde{\mathbf{v}}_{dt}}{\|\tilde{\mathbf{v}}_{dt}\|_1},$$

where $\hat{y}_{dt}^{\text{true}}$, $\hat{y}_{dt}^{\text{cross}}$, and $\hat{y}_{dt}^{\text{diag}}$ are the expected costs predicted with, respectively, the ground truth model, the cross-asset (non-diagonal) model estimated above, and the idiosyncratic (diagonal) model. We also compute associated R^2 values as a performance measure of each model: with $\bar{y} \triangleq \frac{1}{DT}\sum_d \sum_t y_{dt}^{\text{real}}$,

$$R_{\text{true}}^2 \triangleq 1-\frac{\sum_d \sum_t (y_{dt}^{\text{real}} - \hat{y}_{dt}^{\text{true}})^2}{\sum_d \sum_t (y_{dt}^{\text{real}} - \bar{y})^2},\ R_{\text{cross}}^2 \triangleq 1-\frac{\sum_d \sum_t (y_{dt}^{\text{real}} - \hat{y}_{dt}^{\text{cross}})^2}{\sum_d \sum_t (y_{dt}^{\text{real}} - \bar{y})^2},\ R_{\text{diag}}^2 \triangleq 1-\frac{\sum_d \sum_t (y_{dt}^{\text{real}} - \hat{y}_{dt}^{\text{diag}})^2}{\sum_d \sum_t (y_{dt}^{\text{real}} - \bar{y})^2}.$$

From Figure 9 (right), we can visually verify that, in this simulated setup, the estimated model improves the accuracy of prediction compared to the prevalent idiosyncratic (diagonal) impact model. Table 2 shows R^2 values of each model for each intraday segment: while the overall R^2 values are small (even for the ground truth model) due to the high noise level in return realizations, the estimated model works better than the idiosyncratic model.

Intraday segment	R_{true}^2	R_{cross}^2	R_{diag}^2
Beginning of the day (09:30–10:30)	0.1123	0.1184	0.1043
Middle of the day (10:30–15:00)	0.0718	0.0673	0.0611
End of the day (15:00–16:00)	0.0729	0.0683	0.0639
All day (09:30–16:00)	0.1019	0.1023	0.0927

Table 2: R^2 values of the ground truth model, our suggested model, and a separable diagonal impact model on a synthetic dataset.

We expect that the estimation may not work well in some situations. For example, if the realized transactions account for a negligible fraction of the total market volume, their market impact will be insignificant and hardly distinguishable from the noise, and this may result in inaccurate estimates. If the realized transactions are mainly driven by the single-stock level investments, their aggregate impact on the index-fund prices will be relatively smaller than their impact on the individual stock prices, and hence the estimates related to index-fund liquidity (i.e., $\widehat{\nu}_{\text{f},kt}$) will involve larger

[6] The idiosyncratic impact model can be seen as a special case of our estimation model where index-fund investors do not exist (i.e., $\tilde{\mathbf{G}}_{dt}(\nu_t^{\text{diag}}, \mathbf{0})$). For the estimation of the diagonal model, we find the best coefficient $\widehat{\nu}_t^{\text{diag}}$ that minimizes the loss $\mathcal{L}_{\text{id}}$ for each intraday segment (only steps 1 & 2 are performed).

estimation errors than the ones related to single-stock liquidity (i.e., $\widehat{\nu}_{\mathrm{id},t}$). A similar issue may arise when the market liquidity provision is mainly driven by the index-fund investors since the impact on the index fund prices will be relatively small. We observe relatively large estimation errors for the last intraday segment in our numerical demonstration (Figure 9), and this would be partly attributable to the above concern because near the end of the day a larger proportion of liquidity is provided along the index funds.

D. Proofs

D.1. Proof of Proposition 2

We first focus on the case where $\mathbf{v} = \mathbf{W}\mathbf{u}$ in (15). When $\alpha = \beta = 1$, by Woodbury's identity we get

$$\mathbf{G} = \left(\mathbf{\Psi}_{\mathrm{id}} + \mathbf{W}\mathbf{\Psi}_{\mathrm{f}}\mathbf{W}^\top\right)^{-1} = \mathbf{\Psi}_{\mathrm{id}}^{-1} - \mathbf{\Psi}_{\mathrm{id}}^{-1}\mathbf{W}\left(\mathbf{\Psi}_{\mathrm{f}}^{-1} + \mathbf{W}^\top\mathbf{\Psi}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^\top\mathbf{\Psi}_{\mathrm{id}}^{-1}.$$

Consequently,

$$\begin{aligned}\mathbf{W}^\top\mathbf{G}\mathbf{W} &= \mathbf{W}^\top\mathbf{\Psi}_{\mathrm{id}}^{-1}\mathbf{W} - \mathbf{W}^\top\mathbf{\Psi}_{\mathrm{id}}^{-1}\mathbf{W}\left(\mathbf{\Psi}_{\mathrm{f}}^{-1} + \mathbf{W}^\top\mathbf{\Psi}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^\top\mathbf{\Psi}_{\mathrm{id}}^{-1}\mathbf{W} \\ &= \left(\left(\mathbf{W}^\top\mathbf{\Psi}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1} + \mathbf{\Psi}_{\mathrm{f}}\right)^{-1}.\end{aligned}$$

Next, we incorporate the effect of α and β as follows:

$$\mathbf{W}^\top\mathbf{G}\mathbf{W} = \left(\alpha \cdot \left(\mathbf{W}^\top\mathbf{\Psi}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1} + \beta \cdot \mathbf{\Psi}_{\mathrm{f}}\right)^{-1} \longrightarrow \mathbf{\Psi}_{\mathrm{f}}^{-1} \quad \text{as } \alpha \to 0 \text{ and } \beta \to 1.$$

Therefore, for any $\mathbf{u} \in \mathbb{R}^K$,

$$\lim_{\alpha\to 0,\beta\to 1} \bar{\mathcal{C}}\left(\mathbf{v} = \mathbf{W}\mathbf{u}\right) = \lim_{\alpha\to 0,\beta\to 1} \frac{1}{2}\mathbf{u}^\top\mathbf{W}^\top\mathbf{G}\mathbf{W}\mathbf{u} = \frac{1}{2}\mathbf{u}^\top\mathbf{\Psi}_{\mathrm{f}}^{-1}\mathbf{u}.$$

Next we consider the case where $\mathbf{v} \notin \mathrm{span}(\mathbf{w}_1, \cdots, \mathbf{w}_K)$. Let $\mathbf{v} = \mathbf{W}\mathbf{u} + \mathbf{e}$ for some $\mathbf{e} \in \mathbb{R}^N$ such that $\mathbf{W}^\top\mathbf{e} = \mathbf{0}$ and $\mathbf{e} \neq \mathbf{0}$. By the Woodbury matrix identity, we get

$$\mathbf{G} = \alpha^{-1} \cdot \mathbf{\Psi}_{\mathrm{id}}^{-1} - \alpha^{-1} \cdot \mathbf{\Psi}_{\mathrm{id}}^{-1}\mathbf{W}\left(\frac{\alpha}{\beta}\mathbf{\Psi}_{\mathrm{f}}^{-1} + \mathbf{W}^\top\mathbf{\Psi}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^\top\mathbf{\Psi}_{\mathrm{id}}^{-1}.$$

Therefore,

$$\lim_{\alpha\to 0,\beta\to 1} \{\alpha \cdot \mathbf{G}\} = \mathbf{\Psi}_{\mathrm{id}}^{-1} - \mathbf{\Psi}_{\mathrm{id}}^{-1}\mathbf{W}\left(\mathbf{W}^\top\mathbf{\Psi}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^\top\mathbf{\Psi}_{\mathrm{id}}^{-1}.$$

With $\mathbf{r} \triangleq \boldsymbol{\Psi}_{\mathrm{id}}^{-1/2}\mathbf{e}$ and $\mathbf{A} \triangleq \boldsymbol{\Psi}_{\mathrm{id}}^{-1/2}\mathbf{W}$,

$$\mathbf{e}^\top \left(\boldsymbol{\Psi}_{\mathrm{id}}^{-1} - \boldsymbol{\Psi}_{\mathrm{id}}^{-1}\mathbf{W}\left(\mathbf{W}^\top\boldsymbol{\Psi}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^\top\boldsymbol{\Psi}_{\mathrm{id}}^{-1}\right)\mathbf{e} = \mathbf{r}^\top\mathbf{r} - \mathbf{r}^\top\mathbf{A}\left(\mathbf{A}^\top\mathbf{A}\right)^{-1}\mathbf{A}^\top\mathbf{r}.$$

Note that $\mathbf{A}\left(\mathbf{A}^\top\mathbf{A}\right)^{-1}\mathbf{A}^\top\mathbf{r}$ is a projection of $\mathbf{r}$ onto the space spanned by $\mathbf{A}$ (denoted by $\mathrm{span}(\mathbf{A})$). Therefore,

$$\lim_{\alpha\to 0,\beta\to 1}\left\{\alpha\cdot\mathbf{e}^\top\mathbf{G}\mathbf{e}\right\} = 0 \quad \text{if and only if} \quad \mathbf{r}\in\mathrm{span}(\mathbf{A}).$$

If $\mathbf{r}\in\mathrm{span}(\mathbf{A})$, i.e., $\mathbf{r} = \mathbf{A}\mathbf{s}$ for some $\mathbf{s}\in\mathbb{R}^K$, then $\mathbf{e} = \boldsymbol{\Psi}_{\mathrm{id}}^{1/2}\mathbf{r} = \boldsymbol{\Psi}_{\mathrm{id}}^{1/2}\boldsymbol{\Psi}_{\mathrm{id}}^{-1/2}\mathbf{W}\mathbf{s} = \mathbf{W}\mathbf{s}$, and hence $\mathbf{v}\in\mathrm{span}(\mathbf{W})$. Since we are assuming $\mathbf{v}\notin\mathrm{span}(\mathbf{W})$, we have $\mathbf{r}\notin\mathrm{span}(\mathbf{A})$, and hence

$$\lim_{\alpha\to 0,\beta\to 1}\left\{\alpha\cdot\mathbf{e}^\top\mathbf{G}\mathbf{e}\right\} > 0.$$

Furthermore,

$$\begin{aligned}
\mathbf{G}\mathbf{W} &= \alpha^{-1}\cdot\boldsymbol{\Psi}_{\mathrm{id}}^{-1}\mathbf{W} - \alpha^{-1}\cdot\boldsymbol{\Psi}_{\mathrm{id}}^{-1}\mathbf{W}\left(\frac{\alpha}{\beta}\boldsymbol{\Psi}_{\mathrm{f}}^{-1} + \mathbf{W}^\top\boldsymbol{\Psi}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^\top\boldsymbol{\Psi}_{\mathrm{id}}^{-1}\mathbf{W} \\
&= \alpha^{-1}\cdot\boldsymbol{\Psi}_{\mathrm{id}}^{-1}\mathbf{W}\underbrace{\left(\mathbf{I}_K - \left[\frac{\alpha}{\beta}\left(\mathbf{W}^\top\boldsymbol{\Psi}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\boldsymbol{\Psi}_{\mathrm{f}}^{-1} + \mathbf{I}_K\right]^{-1}\right)}_{\longrightarrow\mathbf{O}\text{ as }\alpha\to 0}.
\end{aligned}$$

Therefore,

$$\lim_{\alpha\to 0,\beta\to 1}\left\{\alpha\cdot\mathbf{e}^\top\mathbf{G}\mathbf{W}\mathbf{u}\right\} = 0.$$

To summarize, since $\lim_{\alpha\to 0,\beta\to 1}\left\{\mathbf{u}\mathbf{W}^\top\mathbf{G}\mathbf{W}\mathbf{u}\right\} = \mathbf{u}^\top\boldsymbol{\Psi}_{\mathrm{f}}^{-1}\mathbf{u}$, it follows that

$$\begin{aligned}
&\lim_{\alpha\to 0,\beta\to 1}\left\{\alpha\cdot(\mathbf{W}\mathbf{u}+\mathbf{e})^\top\mathbf{G}(\mathbf{W}\mathbf{u}+\mathbf{e})\right\} \\
=\ &\lim_{\alpha\to 0,\beta\to 1}\left\{\alpha\cdot\mathbf{u}\mathbf{W}^\top\mathbf{G}\mathbf{W}\mathbf{u}\right\} + \lim_{\alpha\to 0,\beta\to 1}\left\{\alpha\cdot 2\mathbf{e}^\top\mathbf{G}\mathbf{W}\mathbf{u}\right\} + \lim_{\alpha\to 0,\beta\to 1}\left\{\alpha\cdot\mathbf{e}^\top\mathbf{G}\mathbf{e}\right\} \\
=\ &0 + 0 + \lim_{\alpha\to 0,\beta\to 1}\left\{\alpha\cdot\mathbf{e}^\top\mathbf{G}\mathbf{e}\right\} > 0.
\end{aligned} \tag{56}$$

It then follows that $\lim_{\alpha\to 0,\beta\to 1}\left\{(\mathbf{W}\mathbf{u}+\mathbf{e})^\top\mathbf{G}(\mathbf{W}\mathbf{u}+\mathbf{e})\right\} = \infty$. ■

D.2. Proof of Proposition 4

Note that

$$
\begin{aligned}
& \mathbf{W}\bar{\mathbf{\Psi}}_{\mathrm{f}}\mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1} - \mathbf{W}\bar{\mathbf{\Psi}}_{\mathrm{f}}\mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\left(\bar{\mathbf{\Psi}}_{\mathrm{f}}^{-1} + \mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1} \\
= & \mathbf{W}\bar{\mathbf{\Psi}}_{\mathrm{f}}\cdot\left(\bar{\mathbf{\Psi}}_{\mathrm{f}}^{-1} + \mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)\cdot\left(\bar{\mathbf{\Psi}}_{\mathrm{f}}^{-1} + \mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1} \\
& -\mathbf{W}\bar{\mathbf{\Psi}}_{\mathrm{f}}\cdot\mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\cdot\left(\bar{\mathbf{\Psi}}_{\mathrm{f}}^{-1} + \mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1} \\
= & \mathbf{W}\bar{\mathbf{\Psi}}_{\mathrm{f}}\cdot\left(\bar{\mathbf{\Psi}}_{\mathrm{f}}^{-1} + \mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W} - \mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)\cdot\left(\bar{\mathbf{\Psi}}_{\mathrm{f}}^{-1} + \mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1} \\
= & \mathbf{W}\left(\bar{\mathbf{\Psi}}_{\mathrm{f}}^{-1} + \mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}.
\end{aligned}
$$

Again, using the Woodbury matrix identity, we get

$$
\begin{aligned}
\mathbf{G}_t^{-1}\left(\sum_{s=1}^{T}\mathbf{G}_s^{-1}\right)^{-1} &= \left(\alpha_t\bar{\mathbf{\Psi}}_{\mathrm{id}} + \beta_t\mathbf{W}\bar{\mathbf{\Psi}}_{\mathrm{f}}\mathbf{W}^{\top}\right)\left(\bar{\mathbf{\Psi}}_{\mathrm{id}} + \mathbf{W}\bar{\mathbf{\Psi}}_{\mathrm{f}}\mathbf{W}^{\top}\right)^{-1} \\
&= \left(\alpha_t\bar{\mathbf{\Psi}}_{\mathrm{id}} + \beta_t\mathbf{W}\bar{\mathbf{\Psi}}_{\mathrm{f}}\mathbf{W}^{\top}\right)\left(\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1} - \bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\left(\bar{\mathbf{\Psi}}_{\mathrm{f}}^{-1} + \mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\right) \\
&= \alpha_t\mathbf{I}_N - \alpha_t\mathbf{W}\left(\bar{\mathbf{\Psi}}_{\mathrm{f}}^{-1} + \mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1} \\
&\quad +\beta_t\mathbf{W}\bar{\mathbf{\Psi}}_{\mathrm{f}}\mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1} - \beta_t\mathbf{W}\bar{\mathbf{\Psi}}_{\mathrm{f}}\mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\left(\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1} + \mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1} \\
&= \alpha_t\mathbf{I}_N + (\beta_t - \alpha_t)\mathbf{W}\underbrace{\left(\bar{\mathbf{\Psi}}_{\mathrm{f}}^{-1} + \mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^{\top}\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}}_{\triangleq\widehat{\mathbf{W}}^{\top}}.
\end{aligned}
$$

■

D.3. Simple Generative Order Flow Model used in §5

We first establish an explicit relationship between intraday variation of natural liquidity and intraday variation of the resulting traded volume, by introducing a stochastic-process generative model for trading volume. The underlying motivation is simple yet intuitive: single-stock and index-fund investors create (stochastic) order flows onto the securities they wish to trade. The arrival intensity of these order flows per type of investor in each time period is proportional to the corresponding trading activity or liquidity provided by this investor type in this time period. This is captured by the profiles α_t and β_t, respectively.

Specifically, we assume that the notional trade volume of stock i in time interval t on day d, DVol_{idt}, is composed of order flows made by single-stock investors $Q_{\mathrm{id},idt}$ and a $|\tilde{w}_{1i}|$ proportion of order flows made by index-fund investors $Q_{\mathrm{f},dt}$. We let $|\tilde{w}_{1i}|$ be dollar-weighted ownership of stock

i in the index fund so that trading one dollar amount of an index fund accumulates $|\tilde{w}_{1i}|$ dollar amount of notional trade volume onto stock i. Each order flow can naturally be decomposed into small transactions:

$$\text{DVol}_{idt} = Q_{\text{id},idt} + |\tilde{w}_{1i}| \cdot Q_{\text{f},dt} = \sum_{j=1}^{N_{\text{id},idt}} q_{\text{id},idt}(j) + |\tilde{w}_{1i}| \cdot \sum_{j=1}^{N_{\text{f},dt}} q_{\text{f},dt}(j), \tag{57}$$

where $N_{\text{id},idt}$ and $q_{\text{id},idt}(j)$ represent the number of transactions and the absolute size of the j^{th} transaction made by single-stock investors in time interval t on day d. For the transactions made by index-fund investors, $N_{\text{f},dt}$ and $q_{\text{f},dt}(j)$ are defined analogously. We treat $N_{\text{id},idt}$, $N_{\text{f},dt}$, $q_{\text{id},idt}(j)$ and $q_{\text{f},dt}(j)$ as random variables that follow particular distribution assumptions.

The order arrival processes for the two investor types are assumed to be Poisson with time-varying rates that are proportional to α_t and β_t:

$$N_{\text{id},idt} \sim \text{Poisson}(\alpha_t \cdot \Lambda) \quad \text{and} \quad N_{\text{f},dt} \sim \text{Poisson}(\beta_t \cdot \Lambda). \tag{58}$$

We further assume that the individual order quantities $q_{\text{id},idt}(j)$'s (and $q_{\text{f},dt}(j)$'s) are all independent and identically distributed with the following moment conditions:

$$\mathsf{E}\left[q_{\text{id},idt}(j)\right] = \bar{q}_{\text{id},i}, \quad \text{Var}\left[q_{\text{id},idt}(j)\right] = c_v^2 \cdot \bar{q}_{\text{id},i}^2, \quad \mathsf{E}\left[q_{\text{f},dt}(j)\right] = \bar{q}_{\text{f}}, \quad \text{Var}\left[q_{\text{id},idt}(j)\right] = c_v^2 \cdot \bar{q}_{\text{f}}^2, \tag{59}$$

where c_v represents a coefficient of variation.

Under the above assumptions, the single-stock investors' order flow $Q_{\text{id},idt}$ is a compound Poisson process with the following mean and variance:

$$\begin{aligned}
\mathsf{E}\left[Q_{\text{id},idt}\right] &= \mathsf{E}\left[N_{\text{id},idt}\right] \cdot \mathsf{E}\left[q_{\text{id},idt}(j)\right] = \alpha_t \cdot \Lambda \cdot \bar{q}_{\text{id},i}, && (60)\\
\text{Var}\left[Q_{\text{id},idt}\right] &= \mathsf{E}\left[\text{Var}\left(Q_{\text{id},idt} | N_{\text{id},idt}\right)\right] + \text{Var}\left[\mathsf{E}\left(Q_{\text{id},idt} | N_{\text{id},idt}\right)\right] && (61)\\
&= \mathsf{E}\left[N_{\text{id},idt} \cdot c_v^2 \cdot \bar{q}_{\text{id},i}^2\right] + \text{Var}\left[N_{\text{id},idt} \cdot \bar{q}_{\text{id},i}\right] && (62)\\
&= \alpha_t \cdot \Lambda \cdot (c_v^2 + 1) \cdot \bar{q}_{\text{id},i}^2. && (63)
\end{aligned}$$

The mean and variance of $Q_{\text{f},dt}$ can be expressed in a similar manner. Summing these flows for each security we get that

$$
\begin{aligned}
\mathsf{E}\left[\mathrm{DVol}_{idt}\right] &= \alpha_t \cdot \Lambda \cdot \bar{q}_{\mathrm{id},i} + \beta_t \cdot \Lambda \cdot |\tilde{w}_{1i}| \cdot \bar{q}_{\mathrm{f}}, && (64)\\
\mathrm{Var}\left[\mathrm{DVol}_{idt}\right] &= \alpha_t \cdot \Lambda \cdot (1+c_v^2) \cdot \bar{q}_{\mathrm{id},i}^2 + \beta_t \cdot \Lambda \cdot |\tilde{w}_{1i}|^2 \cdot (1+c_v^2) \cdot \bar{q}_{\mathrm{f}}^2, && (65)\\
\mathrm{Cov}\left[\mathrm{DVol}_{idt}, \mathrm{DVol}_{jdt}\right] &= \beta_t \cdot \Lambda \cdot |\tilde{w}_{1i}| \cdot |\tilde{w}_{1j}| \cdot (1+c_v^2) \cdot \bar{q}_{\mathrm{f}}^2. && (66)
\end{aligned}
$$

The common order flow $Q_{\mathrm{f},dt}$ made by index-fund investors results in a positive correlation between stocks represented in the index.

Define θ_i to be the proportion of daily traded volume generated by index-fund investors out of the total daily traded volume of stock i:

$$
\theta_i \triangleq \frac{\sum_{t=1}^{T} \mathsf{E}\left[|\tilde{w}_{1i}| \cdot Q_{\mathrm{f},dt}\right]}{\sum_{t=1}^{T} \mathsf{E}\left[\mathrm{DVol}_{idt}\right]} = \frac{|\tilde{w}_{1i}| \cdot \bar{q}_{\mathrm{f}}}{\bar{q}_{\mathrm{id},i} + |\tilde{w}_{1i}| \cdot \bar{q}_{\mathrm{f}}}. \tag{67}
$$

The intraday traded volume profile $\mathrm{VolAlloc}_{it}$ and the pairwise correlation Correl_{ijt}, defined in (1) and (2), can be simply expressed with θ_i and θ_j:

$$
\begin{aligned}
\mathrm{VolAlloc}_{it} &\equiv \frac{\mathsf{E}\left[\mathrm{DVol}_{idt}\right]}{\sum_{s=1}^{T} \mathsf{E}\left[\mathrm{DVol}_{ids}\right]} = \alpha_t \cdot (1-\theta_i) + \beta_t \cdot \theta_i, && (68)\\
\mathrm{Correl}_{ijt} &\equiv \frac{\mathrm{Cov}\left[\mathrm{DVol}_{idt}, \mathrm{DVol}_{jdt}\right]}{\sqrt{\mathrm{Var}\left[\mathrm{DVol}_{idt}\right]} \cdot \sqrt{\mathrm{Var}\left[\mathrm{DVol}_{jdt}\right]}} && (69)\\
&= \frac{\beta_t \cdot \theta_i \cdot \theta_j}{\sqrt{\alpha_t \cdot (1-\theta_i)^2 + \beta_t \cdot \theta_i^2} \cdot \sqrt{\alpha_t \cdot (1-\theta_j)^2 + \beta_t \cdot \theta_j^2}}. && (70)
\end{aligned}
$$

If we further assume that the proportions θ_i are the same across all securities,

$$
\theta \equiv \theta_1 = \theta_2 = \cdots = \theta_N, \tag{71}
$$

then $\mathrm{VolAlloc}_{it}$ is the same for all stocks i and Correl_{ijt} is identical across all pairs of stocks, i, j, as given in (28)–(29).

D.4. Proofs for §5.2

D.4.1. Proof of Proposition 5

Note that

$$\begin{aligned}\Upsilon(\mathbf{x}_0) &= \frac{\sum_{t=1}^{T}(\alpha_t\cdot(1-\theta)+\beta_t\cdot\theta)^2\cdot\mathbf{x}_0^\top\left(\alpha_t\bar{\boldsymbol{\Psi}}_{\mathrm{id}}+\beta_t\mathbf{W}\bar{\boldsymbol{\Psi}}_{\mathrm{f}}\mathbf{W}^\top\right)^{-1}\mathbf{x}_0}{\mathbf{x}_0^\top\left(\bar{\boldsymbol{\Psi}}_{\mathrm{id}}+\mathbf{W}\bar{\boldsymbol{\Psi}}_{\mathrm{f}}\mathbf{W}^\top\right)^{-1}\mathbf{x}_0}\\ &= \sum_{t=1}^{T}\alpha_t\cdot(1+\theta\cdot(\gamma_t-1))^2\cdot\frac{\mathbf{x}_0^\top\left(\bar{\boldsymbol{\Psi}}_{\mathrm{id}}+\gamma_t\mathbf{W}\bar{\boldsymbol{\Psi}}_{\mathrm{f}}\mathbf{W}^\top\right)^{-1}\mathbf{x}_0}{\mathbf{x}_0^\top\left(\bar{\boldsymbol{\Psi}}_{\mathrm{id}}+\mathbf{W}\bar{\boldsymbol{\Psi}}_{\mathrm{f}}\mathbf{W}^\top\right)^{-1}\mathbf{x}_0}.\end{aligned}$$

By Woodbury's matrix identity,

$$\begin{aligned}&\frac{\mathbf{x}_0^\top\left(\bar{\boldsymbol{\Psi}}_{\mathrm{id}}+\gamma_t\mathbf{W}\bar{\boldsymbol{\Psi}}_{\mathrm{f}}\mathbf{W}^\top\right)^{-1}\mathbf{x}_0}{\mathbf{x}_0^\top\left(\bar{\boldsymbol{\Psi}}_{\mathrm{id}}+\mathbf{W}\bar{\boldsymbol{\Psi}}_{\mathrm{f}}\mathbf{W}^\top\right)^{-1}\mathbf{x}_0}\\ =\;&\frac{\mathbf{x}_0^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0-\mathbf{x}_0^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\left(\gamma_t^{-1}\bar{\boldsymbol{\Psi}}_{\mathrm{f}}^{-1}+\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0}{\mathbf{x}_0^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0-\mathbf{x}_0^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\left(\bar{\boldsymbol{\Psi}}_{\mathrm{f}}^{-1}+\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0}\\ =\;&1+\frac{\left(\mathbf{x}_0^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\left(\bar{\boldsymbol{\Psi}}_{\mathrm{f}}^{-1}+\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0\right)-\left(\mathbf{x}_0^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\left(\gamma_t^{-1}\bar{\boldsymbol{\Psi}}_{\mathrm{f}}^{-1}+\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0\right)}{\mathbf{x}_0^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0-\mathbf{x}_0^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\left(\bar{\boldsymbol{\Psi}}_{\mathrm{f}}^{-1}+\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0}\\ =\;&1+\frac{\mathbf{x}_0^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\left(\left(\bar{\boldsymbol{\Psi}}_{\mathrm{f}}^{-1}+\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}-\left(\gamma_t^{-1}\bar{\boldsymbol{\Psi}}_{\mathrm{f}}^{-1}+\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\right)\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0}{\mathbf{x}_0^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0-\mathbf{x}_0^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\left(\bar{\boldsymbol{\Psi}}_{\mathrm{f}}^{-1}+\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0}\\ =\;&1+\frac{\mathbf{x}_0^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\left(\bar{\boldsymbol{\Psi}}_{\mathrm{f}}^{-1}+\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\cdot\left(\gamma_t^{-1}-1\right)\bar{\boldsymbol{\Psi}}_{\mathrm{f}}^{-1}\cdot\left(\gamma_t^{-1}\bar{\boldsymbol{\Psi}}_{\mathrm{f}}^{-1}+\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0}{\mathbf{x}_0^\top\left(\bar{\boldsymbol{\Psi}}_{\mathrm{id}}+\mathbf{W}\bar{\boldsymbol{\Psi}}_{\mathrm{f}}\mathbf{W}^\top\right)^{-1}\mathbf{x}_0}\\ =\;&1+(1-\gamma_t)\cdot\frac{\mathbf{x}_0^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\left(\bar{\boldsymbol{\Psi}}_{\mathrm{f}}^{-1}+\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\left(\mathbf{I}_K+\gamma_t\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\bar{\boldsymbol{\Psi}}_{\mathrm{f}}\right)^{-1}\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0}{\mathbf{x}_0^\top\left(\bar{\boldsymbol{\Psi}}_{\mathrm{id}}+\mathbf{W}\bar{\boldsymbol{\Psi}}_{\mathrm{f}}\mathbf{W}^\top\right)^{-1}\mathbf{x}_0}.\end{aligned}$$

When $K=1$, we get

$$\left(\mathbf{I}_K+\gamma_t\mathbf{W}^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\bar{\boldsymbol{\Psi}}_{\mathrm{f}}\right)^{-1}=\left(1+\gamma_t\mathbf{w}_1^\top\bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{w}_1\bar{\psi}_{\mathrm{f},1}\right)^{-1}=\frac{1}{1+\gamma_t\cdot\eta_1}.$$

Consequently,

$$
\begin{aligned}
&\frac{\mathbf{x}_0^\top \left(\bar{\mathbf{\Psi}}_{\mathrm{id}} + \gamma_t \mathbf{W}\bar{\mathbf{\Psi}}_{\mathrm{f}}\mathbf{W}^\top\right)^{-1}\mathbf{x}_0}{\mathbf{x}_0^\top \left(\bar{\mathbf{\Psi}}_{\mathrm{id}} + \mathbf{W}\bar{\mathbf{\Psi}}_{\mathrm{f}}\mathbf{W}^\top\right)^{-1}\mathbf{x}_0} \\
&= 1 + \frac{1-\gamma_t}{1+\eta_1\cdot\gamma_t}\cdot \frac{\mathbf{x}_0^\top \bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\left(\bar{\mathbf{\Psi}}_{\mathrm{f}}^{-1} + \mathbf{W}^\top\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^\top\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0}{\mathbf{x}_0^\top \bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0 - \mathbf{x}_0^\top \bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\left(\bar{\mathbf{\Psi}}_{\mathrm{f}}^{-1} + \mathbf{W}^\top\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^\top\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0} \\
&= 1 + \frac{1-\gamma_t}{1+\eta_1\cdot\gamma_t}\cdot \left(\frac{\mathbf{x}_0^\top \bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0}{\mathbf{x}_0^\top \bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\left(\bar{\mathbf{\Psi}}_{\mathrm{f}}^{-1} + \mathbf{W}^\top\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{W}\right)^{-1}\mathbf{W}^\top\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0} - 1\right)^{-1} \\
&= 1 + \frac{1-\gamma_t}{1+\eta_1\cdot\gamma_t}\cdot \left(\frac{\mathbf{x}_0^\top \bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0}{\mathbf{x}_0^\top \bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{w}_1\cdot\frac{\bar{\psi}_{\mathrm{f},1}}{1+\bar{\psi}_{\mathrm{f},1}\mathbf{w}_1^\top\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{w}_1}\cdot\mathbf{w}_1^\top\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0} - 1\right)^{-1} \\
&= 1 + \frac{1-\gamma_t}{1+\eta_1\cdot\gamma_t}\cdot \left(\frac{\mathbf{x}_0^\top \bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0}{\left(\mathbf{w}_1^\top\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0\right)^2}\cdot\frac{1+\eta_1}{\bar{\psi}_{\mathrm{f},1}} - 1\right)^{-1}.
\end{aligned}
$$

To simplify notation, define

$$
f(x) \triangleq \left(\frac{\mathbf{x}_0^\top \bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0}{\left(\mathbf{w}_1^\top\bar{\mathbf{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0\right)^2}\cdot\frac{1+\eta_1}{\bar{\psi}_{\mathrm{f},1}} - 1\right)^{-1}.
$$

Then,

$$
\Upsilon(\mathbf{x}_0) = \sum_{t=1}^{T}\alpha_t\cdot(1+\theta\cdot(\gamma_t-1))^2\cdot\left(1+\frac{1-\gamma_t}{1+\eta_1\cdot\gamma_t}\cdot f(\mathbf{x}_0)\right).
$$

Note that

$$
\begin{aligned}
\sum_{t=1}^{T}\alpha_t\cdot(1+\theta\cdot(\gamma_t-1))^2 &= \sum_{t=1}^{T}\alpha_t\cdot\left(1+2\theta\cdot(\gamma_t-1)+\theta^2\cdot(\gamma_t^2-2\gamma_t+1)\right) \\
&= \sum_{t=1}^{T}\alpha_t + 2\theta\cdot(\beta_t-\alpha_t)+\theta^2\cdot\left(\frac{\beta_t^2}{\alpha_t}-2\beta_t+\alpha_t\right) \\
&= 1+\theta^2\cdot\left(\sum_{t=1}^{T}\frac{\beta_t^2}{\alpha_t}-1\right).
\end{aligned}
$$

As a result,

$$\Upsilon(\mathbf{x}_0) = 1 + \theta^2 \cdot \left(\sum_{t=1}^{T} \frac{\beta_t^2}{\alpha_t} - 1\right) + \underbrace{\left(\sum_{t=1}^{T} \frac{\alpha_t \cdot (1-\theta\cdot(1-\gamma_t))^2\,(1-\gamma_t)}{1+\eta_1\cdot\gamma_t}\right)}_{\triangleq \Delta} \times f(\mathbf{x}_0).$$

■

D.5. Proof of Remarks 1 – 3

Maximum/minimum cost ratio. Note that $f(\mathbf{x}_0)$ is a decreasing function of $\frac{\mathbf{x}_0^\top \bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0}{(\mathbf{w}_1^\top \bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0)^2}$, and

$$\begin{aligned}\min_{\mathbf{x}_0\in\mathbb{R}^N} \frac{\mathbf{x}_0^\top \bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0}{\left(\mathbf{w}_1^\top \bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0\right)^2} &= \left(\max_{\mathbf{x}_0\in\mathbb{R}^N} \frac{\left(\mathbf{w}_1^\top \bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0\right)^2}{\mathbf{x}_0^\top \bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0}\right)^{-1} = \left(\max_{\mathbf{y}\in\mathbb{R}^N} \frac{\left(\mathbf{w}_1^\top \bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1/2}\mathbf{y}\right)^2}{\mathbf{y}^\top\mathbf{y}}\right)^{-1} \\ &= \left(\mathbf{w}_1^\top \bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{w}_1\right)^{-1} = \frac{\bar{\psi}_{\mathrm{f},1}}{\eta_1}.\end{aligned}$$

The above value is obtained at $\mathbf{x}_0 = \mathbf{w}_1$. Therefore,

$$\max_{\mathbf{x}_0\in\mathbb{R}^N} f(\mathbf{x}_0) = f(\mathbf{x}_0 = \mathbf{w}_1) = \left(\frac{\bar{\psi}_{\mathrm{f},1}}{\eta_1}\cdot\frac{1+\eta_1}{\bar{\psi}_{\mathrm{f},1}} - 1\right)^{-1} = \eta_1.$$

On the other hand, since $\min_{\mathbf{x}_0\in\mathbb{R}^N} \frac{(\mathbf{w}_1^\top \bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0)^2}{\mathbf{x}_0^\top \bar{\boldsymbol{\Psi}}_{\mathrm{id}}^{-1}\mathbf{x}_0} = 0$ at $\mathbf{x}_0 = \mathbf{w}_1^\perp$, it follows that

$$\min_{\mathbf{x}_0\in\mathbb{R}^N} f(\mathbf{x}_0) = f(\mathbf{x}_0 = \mathbf{w}_1^\perp) = 0.$$

Combining these two results, we have that

$$\begin{aligned}\max_{\mathbf{x}_0\in\mathbb{R}^N} \Upsilon(\mathbf{x}_0) &= 1+\theta^2\cdot\left(\sum_{t=1}^{T}\frac{\beta_t^2}{\alpha_t}-1\right) + \max_{\mathbf{x}_0\in\mathbb{R}^N}\{\Delta\cdot f(\mathbf{x}_0)\} = \max\{\Upsilon_{\mathrm{market}}, \Upsilon_{\mathrm{orth}}\} \\ \min_{\mathbf{x}_0\in\mathbb{R}^N} \Upsilon(\mathbf{x}_0) &= 1+\theta^2\cdot\left(\sum_{t=1}^{T}\frac{\beta_t^2}{\alpha_t}-1\right) + \min_{\mathbf{x}_0\in\mathbb{R}^N}\{\Delta\cdot f(\mathbf{x}_0)\} = \min\{\Upsilon_{\mathrm{market}}, \Upsilon_{\mathrm{orth}}\},\end{aligned}$$

and, trivially, $\Upsilon_{\mathrm{market}} \geq \Upsilon_{\mathrm{orth}}$ if and only if $\Delta \geq 0$.

Sign of Δ with respect to θ. Note that $\Delta(\theta)$ is a quadratic function of θ. It suffices to show that $\Delta(\theta=0)\geq 0$ and $\Delta(\theta=1)\leq 0$. Note that for an arbitrary function $h(\cdot)$,

$$\begin{cases}\text{if } h(\cdot) \text{ is non-decreasing}, & h(\gamma)\cdot(1-\gamma) \leq h(1)\cdot(1-\gamma), \quad \forall\gamma \\ \text{if } h(\cdot) \text{ is non-increasing}, & h(\gamma)\cdot(1-\gamma) \geq h(1)\cdot(1-\gamma), \quad \forall\gamma\end{cases}. \tag{72}$$

In the case of $\theta = 0$, by setting $h(\gamma_t) \triangleq \frac{1}{1+\eta_1 \cdot \gamma_t}$ which is a non-increasing function, we get

$$\Delta(\theta = 0) = \sum_{t=1}^{T} \frac{\alpha_t \cdot (1-\gamma_t)}{1+\eta_1 \cdot \gamma_t} \overset{(72)}{\geq} \sum_{t=1}^{T} \frac{\alpha_t \cdot (1-\gamma_t)}{1+\eta_1} = (1+\eta_1)^{-1} \sum_{t=1}^{T} (\alpha_t - \beta_t) = 0.$$

In the case of $\theta = 1$, by setting $h(\gamma_t) \triangleq \frac{\gamma_t^2}{1+\eta_1 \cdot \gamma_t}$, which is a non-decreasing function, we get

$$\Delta(\theta = 1) = \sum_{t=1}^{T} \frac{\alpha_t \cdot \gamma_t^2 (1-\gamma_t)}{1+\eta_1 \cdot \gamma_t} \overset{(72)}{\leq} \sum_{t=1}^{T} \frac{\alpha_t \cdot (1-\gamma_t)}{1+\eta_1} = 0.$$

Change of Υ_{market} with respect to η_1. Note that

$$\begin{aligned}
\frac{\partial \Upsilon_{\text{market}}}{\partial \eta_1} &= \frac{\partial}{\partial \eta_1} \left(\eta_1 \cdot \Delta(\eta_1) \right) \\
&= \frac{\partial}{\partial \eta_1} \left(\sum_{t=1}^{T} \frac{\alpha_t \cdot (1-\theta \cdot (1-\gamma_t))^2 (1-\gamma_t)}{\eta_1^{-1} + \gamma_t} \right) \\
&= \sum_{t=1}^{T} \frac{\alpha_t \cdot (1-\theta \cdot (1-\gamma_t))^2 (1-\gamma_t)}{\eta_1^2 \cdot (\eta_1^{-1} + \gamma_t)^2} \\
&= \frac{\theta^2}{\eta_1^2} \cdot \sum_{t=1}^{T} \alpha_t \cdot \left(1 + \frac{\theta^{-1} - 1 - \eta_1^{-1}}{\eta_1^{-1} + \gamma_t} \right)^2 (1-\gamma_t).
\end{aligned}$$

Set $h(\gamma_t) \triangleq \left(1 + \frac{\theta^{-1} - 1 - \eta_1^{-1}}{\eta_1^{-1} + \gamma_t} \right)^2$. If $\eta_1 \leq \frac{\theta}{1-\theta}$, then $\theta^{-1} - 1 - \eta_1^{-1} \leq 0$, and hence $h(\cdot)$ is non-decreasing. Therefore,

$$\begin{aligned}
\frac{\partial \Upsilon_{\text{market}}}{\partial \eta_1} &= \frac{\theta^2}{\eta_1^2} \cdot \sum_{t=1}^{T} \alpha_t \cdot \left(1 + \frac{\theta^{-1} - 1 - \eta_1^{-1}}{\eta_1^{-1} + \gamma_t} \right)^2 (1-\gamma_t) \\
&\overset{(72)}{\leq} \frac{\theta^2}{\eta_1^2} \cdot \sum_{t=1}^{T} \alpha_t \cdot \left(1 + \frac{\theta^{-1} - 1 - \eta_1^{-1}}{\eta_1^{-1} + 1} \right)^2 (1-\gamma_t) \\
&= \frac{\theta^2}{\eta_1^2} \cdot \left(1 + \frac{\theta^{-1} - 1 - \eta_1^{-1}}{\eta_1^{-1} + 1} \right)^2 \cdot \sum_{t=1}^{T} \alpha_t (1-\gamma_t) \\
&= \frac{\theta^2}{\eta_1^2} \cdot \left(1 + \frac{\theta^{-1} - 1 - \eta_1^{-1}}{\eta_1^{-1} + 1} \right)^2 \cdot \sum_{t=1}^{T} (\alpha_t - \beta_t) \\
&= 0.
\end{aligned}$$

If $\eta_1 \geq \frac{\theta}{1-\theta}$, then $h(\cdot)$ is non-increasing, and hence the sign of the inequality reverses. Therefore,

$$\frac{\partial \Upsilon_{\text{market}}}{\partial \eta_1} \leq 0 \quad \text{if } \eta_1 \leq \frac{\theta}{1-\theta}, \quad \text{and} \quad \frac{\partial \Upsilon_{\text{market}}}{\partial \eta_1} \geq 0 \quad \text{if } \eta_1 \geq \frac{\theta}{1-\theta}.$$

■

When $\eta_1 = 0$,

$$\Upsilon_{\text{market}}(\eta_1 = 0) = 1 + \theta^2 \cdot \left(\sum_{t=1}^{T} \frac{\beta_t^2}{\alpha_t} - 1 \right).$$

Note that $\Upsilon_{\text{market}}(\eta_1 = 0) = \Upsilon_{\text{orth}}$. Since $\Upsilon_{\text{market}}(\eta_1)$ is decreasing in $[0, \frac{\theta}{1-\theta}]$, this completes proof of (40). When $\eta_1 = \frac{\theta}{1-\theta}$, since $\frac{(1-\theta+\theta\cdot\gamma_t)^2}{1+\eta_1\cdot\gamma_t} = (1-\theta)\cdot(1-\theta+\theta\cdot\gamma_t)$, it follows that

$$\begin{aligned}
\Upsilon_{\text{market}}(\eta_1 = \frac{\theta}{1-\theta}) &= 1 + \theta^2 \cdot \left(\sum_{t=1}^{T} \frac{\beta_t^2}{\alpha_t} - 1 \right) + \frac{\theta}{1-\theta} \cdot (1-\theta) \cdot \sum_{t=1}^{T} \alpha_t \cdot (1 - \theta \cdot (1-\gamma_t))\,(1-\gamma_t) \\
&= 1 + \theta^2 \cdot \left(\sum_{t=1}^{T} \frac{\beta_t^2}{\alpha_t} - 1 \right) - \theta^2 \cdot \left(\sum_{t=1}^{T} \frac{\beta_t^2}{\alpha_t} - 1 \right) \\
&= 1.
\end{aligned}$$

As $\eta_1 \to \infty$,

$$\begin{aligned}
\lim_{\eta_1 \to \infty} \Upsilon_{\text{market}} &= 1 + \theta^2 \cdot \left(\sum_{t=1}^{T} \frac{\beta_t^2}{\alpha_t} - 1 \right) + \lim_{\eta_1 \to \infty} (\eta_1 \cdot \Delta(\eta_1)) \\
&= 1 + \theta^2 \cdot \left(\sum_{t=1}^{T} \frac{\beta_t^2}{\alpha_t} - 1 \right) + \sum_{t=1}^{T} \frac{\alpha_t \cdot (1 - \theta \cdot (1-\gamma_t))^2\,(1-\gamma_t)}{\gamma_t} \\
&= 1 + \theta^2 \cdot \left(\sum_{t=1}^{T} \frac{\beta_t^2}{\alpha_t} - 1 \right) + (1-\theta)^2 \cdot \left(\sum_{t=1}^{T} \frac{\alpha_t^2}{\beta_t} \right) - 1 + 2\theta - \theta^2 \cdot \left(\sum_{t=1}^{T} \frac{\beta_t^2}{\alpha_t} \right) \\
&= 1 + (1-\theta)^2 \cdot \left(\sum_{t=1}^{T} \frac{\alpha_t^2}{\beta_t} - 1 \right).
\end{aligned}$$

■

Cost ratio of single-stock trading. Note that

$$\begin{aligned}
f(\mathbf{e}_i) &= \left(\frac{\mathbf{e}_i^\top \bar{\boldsymbol{\Psi}}_{\text{id}}^{-1} \mathbf{e}_i}{\left(\mathbf{w}_1^\top \bar{\boldsymbol{\Psi}}_{\text{id}}^{-1} \mathbf{e}_i\right)^2} \cdot \frac{1+\eta_1}{\bar{\psi}_{\text{f},1}} - 1 \right)^{-1} = \left(\frac{\bar{\psi}_{\text{id},i}^{-1}}{\left(w_{1i} \cdot \bar{\psi}_{\text{id},i}^{-1}\right)^2} \cdot \frac{1+\eta_1}{\bar{\psi}_{\text{f},1}} - 1 \right)^{-1} \\
&= \left(\frac{1+\eta_1}{\eta_{1,i}} - 1 \right)^{-1} = \frac{\eta_{1,i}}{1 + \eta_1 - \eta_{1,i}}.
\end{aligned}$$

Also note that

$$\frac{\eta_{1,i}}{1+\eta_1-\eta_{1,i}} \geq \frac{\eta_{1,j}}{1+\eta_1-\eta_{1,j}} \quad \text{if and only if} \quad \frac{w_{1i}^2}{\bar{\psi}_{\text{id},i}} \geq \frac{w_{1j}^2}{\bar{\psi}_{\text{id},j}}.$$

The results immediately follow from (34). ■