SIAM J. Financial Math.
Vol. 11, No. 2, pp. 327–357

The Multivariate Kyle Model: More is Different*

Luis Carlos Garcia del Molino†, Iacopo Mastromatteo†, Michael Benzaquen‡, and Jean-Philippe Bouchaud†

Abstract. We reconsider the multivariate Kyle model in a risk-neutral setting with a single, perfectly informed rational insider and a rational competitive market maker, setting the price of n securities. We prove the unicity of a symmetric, positive definite solution for the impact matrix and provide insights on its interpretation. We explore its implications from the perspective of empirical market microstructure and argue that it provides a sensible inference procedure to cure some pathologies encountered in recent attempts to calibrate cross-impact matrices. As an illustration, we determine the empirical cross-impact matrix of US Treasuries and compare the results with recent alternative calibration methods.

Key words. microstructure, impact, multivariate, price formation

AMS subject classifications. 15A24, 62P20, 91A40, 91B26

DOI. 10.1137/18M1231997

1. Introduction.

Understanding market impact—the mechanism through which trades tend to push prices—is with no doubt a venture of paramount importance. From the theoretical point of view, market impact lies at the very heart of price formation in financial markets. From the practitioner's perspective, market impact is often at the origin of nonnegligible trading costs that need to be controlled to optimize execution strategies. In the past decades, with the notable exception of [4], from which our theoretical results stem, most of the literature has focused on the price impact of single products (for a recent review, see [3]), with no regard to interasset interactions. However, many market participants trade large portfolios that combine hundreds or thousands of assets. Thus addressing the matter of interasset price impact, coined *cross-impact*, is of great interest both fundamentally and practically. Recent empirical studies show evidence of significant cross-impact effects in stock markets [7, 13, 16, 2]. A recurrent issue is that of empirical noise when it comes to large matrix estimation. It is thus important to continue to search for good statistical priors in order to help "clean" these large dimensional estimators.

What should one expect from theoretical economics on this matter? Which empirical observations should be considered unusual? Here we focus on how to make sense of the empirical observations accumulated over the past few years in the most orthodox setting: the classic Kyle model [9], extended to a multiasset framework. Our aim is to gain insight into

*Received by the editors December 12, 2018; accepted for publication (in revised form) January 28, 2020; published electronically April 2, 2020.
https://doi.org/10.1137/18M1231997

†Capital Fund Management, 23 rue de l'Université, 75007, Paris, France (garciadelmolino@gmail.com, Iacopo.Mastromatteo@cfm.fr, Jean-Philippe.BOUCHAUD@cfm.fr).

‡Capital Fund Management, 23 rue de l'Université, 75007, Paris, France and Ladhyx, UMR CNRS 7646, Ecole polytechnique, 91128 Palaiseau Cedex, France (michael.benzaquen@ladhyx.polytechnique.fr).

(i) how information is diffused into prices (cross-sectionally) within the Kyle setting and (ii) how one can use such results to regularize the very noisy regressions that arise in empirical cross-impact analysis. The multivariate Kyle model was first considered in [4, 5], in a very general setting with n assets and m partially informed traders. Here we revisit the problem with the issue of empirical calibration in mind and we focus on the particular case $m = 1$. The equilibrium solution provides an explicit recipe to infer cross impact from (cleaned) order flow and return correlation matrices, which we compare to other natural recipes, such as maximum likelihood estimators (MLE) [2] or the recently proposed "EigenLiquidity" Model (ELM) [12].

The paper is organized as follows. In sections 2 and 3 we introduce the multivariate Kyle model and provide the equilibrium strategies, mostly building upon the work in [4] but also providing new results. In section 4 we study in detail the mathematical properties of the solution. In section 5, we introduce an impact estimator based on the equilibrium strategy of the market maker (MM) and compare it to other usual estimators both from the theoretical and empirical points of view. Throughout the text we present several examples intended to provide intuition behind the main results. In particular it is often interesting to confront the results of the multivariate and univariate models.

In all of the following, bold uppercase symbols denote matrices, bold lowercases denote vectors, and light lowercases denote scalars.

2. The multivariate Kyle model.

In this section, we present in detail the multivariate Kyle setting, the special case of [4] with $m = 1$, and define the observables that would allow one to calibrate the model using empirical observables.

2.1. The model.

Consider a single-period economy where three representative agents trade n instruments. The agents are an informed trader (IT) who has perfect information about the future prices $\boldsymbol{v}$, a noise trader (NT) that trades in the absence of any information due to exogenous reasons, and a competitive market maker that has the role of enforcing price efficiency. The dynamics of the model is set by the following rules.

1. A fundamental price $\boldsymbol{v}$ is sampled from a Gaussian distribution $\boldsymbol{v} \sim \mathcal{N}(\boldsymbol{p}_0, \boldsymbol{\Sigma}_0)$, where $\boldsymbol{\Sigma}_0$ is SPD. Only the IT knows the value of $\boldsymbol{v}$ in advance, while $\boldsymbol{p}_0$ and $\boldsymbol{\Sigma}_0$ are common knowledge. We denote the price deviation from its mean as $\Delta\boldsymbol{v} := \boldsymbol{v} - \boldsymbol{p}_0$.
2. The IT and NT place simultaneously their orders of sizes $\boldsymbol{x}$ an $\boldsymbol{u}$, respectively. The bids of the NT $\boldsymbol{u}$ are sampled from a Gaussian distribution $\boldsymbol{u} \sim \mathcal{N}(0, \boldsymbol{\Omega})$ independent of the fundamental price, where $\boldsymbol{\Omega}$ is an invertible matrix.
3. The MM clears the excess demand $\boldsymbol{y}$ at a clearing price $\boldsymbol{p}$ based on the total observed order imbalance $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{u}$, which allows him/her to form the best estimation of the fundamental prices. These fundamental prices are then revealed.

The quantity $\boldsymbol{x}$ requested by the *risk-neutral* IT is such that he maximizes the expectation of his utility function:

$$\mathcal{U}_{IT}(\boldsymbol{x}, \boldsymbol{p}) = \boldsymbol{x}^\top(\boldsymbol{v} - \boldsymbol{p})\,. \tag{2.1}$$

Note that $\mathcal{U}_{IT}$ does not contain any risk penalty. While the introduction of such a penalty would affect some of the conclusions below, we decide to leave this interesting issue for future work. Consistent with the assumption that market making is competitive, the MM sets a price that matches in expectation $\boldsymbol{v}$ given the available public information, namely, the total order imbalance $\boldsymbol{y}$:

$$\boldsymbol{p} = \mathbb{E}[\boldsymbol{v}|\boldsymbol{y}]\,, \tag{2.2}$$

where $\mathbb{E}[\cdot]$ denotes the average with respect to the distribution of $\boldsymbol{v}$ and $\boldsymbol{u}$.

Clearly, the above setting is highly stylized and, on many counts, unrealistic.[1] Still, this model is able to capture some of the essential ingredients of a reasonable price-formation process: the information owned by ITs gets encoded into a trading order imbalance $\boldsymbol{x}$ polluted by a noise $\boldsymbol{u}$. In a competitive regime, MMs are expected to decode the information contained in the total order imbalance $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{u}$, in order to provide the best possible prediction of the fundamental price $\boldsymbol{v}$. This mechanically induces market impact: because the order imbalance $\boldsymbol{y}$ is correlated with the fundamental price, the traded price $\boldsymbol{p}$ will also display a correlation with the order imbalance $\boldsymbol{y}$, thus providing a sensible measure of market impact.

2.2. Observables. In order to gain insight into the implications of this model and in order to make testable predictions, one is required to provide some metrics that can be compared against market data. Luckily enough, the Gaussian nature of the setup allows one to only consider first and second order statistics (means and covariances) of prices and volumes in order to completely characterize the behavior of the model. Naturally, the fundamental parameters defining the model $(\boldsymbol{\Omega}, \boldsymbol{\Sigma}_0)$ are not directly observable and need to be inferred from the statistics of trades prices $\boldsymbol{p}$, and of volumes $\boldsymbol{y}$, which are the only physical observables of the model.

Prices. Due to the price efficiency condition (see (2.2)), the average traded price is equal to the average of the fundamental price, itself equal (for consistency) to the initial price:

$$\mathbb{E}[\boldsymbol{p}] = \mathbb{E}[\boldsymbol{v}] = \boldsymbol{p}_0\,. \tag{2.3}$$

However, the covariance of the traded price and that of the fundamental price have no reason to coincide:

$$\boldsymbol{\Sigma} = \mathbb{C}[\boldsymbol{p},\boldsymbol{p}] = \mathbb{E}[(\boldsymbol{p}-\boldsymbol{p}_0)(\boldsymbol{p}-\boldsymbol{p}_0)^\top]\,, \tag{2.4}$$

$$\boldsymbol{\Sigma}_0 = \mathbb{C}[\boldsymbol{v},\boldsymbol{v}] = \mathbb{E}[(\boldsymbol{v}-\boldsymbol{p}_0)(\boldsymbol{v}-\boldsymbol{p}_0)^\top]\,. \tag{2.5}$$

Further down we show that, at equilibrium, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}_0$ are exactly proportional (see (3.13)).

Volumes. Due to the uninformed nature of the NT, the average order imbalance is fixed by the bias introduced by the IT, so that

$$\mathbb{E}[\boldsymbol{y}] = \mathbb{E}[\boldsymbol{x}] := \boldsymbol{y}_0\,. \tag{2.6}$$

The relation between the portion of volume covariance due to the NT and the one due to the IT is more subtle and is given by

$$\boldsymbol{\Omega}^{\mathrm{d}} = \mathbb{C}[\boldsymbol{y},\boldsymbol{y}] = \mathbb{C}[\boldsymbol{x},\boldsymbol{x}] + \boldsymbol{\Omega}\;, \tag{2.7}$$

[1] For example, the order flow in a Kyle setting has no temporal auto-correlations, whereas it is well known that the empirical order flow has a long memory. See [3] and [2] for a recent discussion in the multivariate context.

where we have introduced what we have coined the *dressed volume covariance* $\boldsymbol{\Omega}^{\mathrm{d}}$, which is the physically observable quantity. Since the IT places the bid without knowing $\boldsymbol{u}$ (he only knows the *statistics* of $\boldsymbol{u}$), $\mathbb{C}[\boldsymbol{x}, \boldsymbol{u}] = 0$ and therefore the *bare* volume covariance (i.e., not dressed by the noise contribution $\boldsymbol{\Omega}$) is $\mathbb{C}[\boldsymbol{x}, \boldsymbol{x}]$.

Response. Ultimately, we are interested in characterizing the expected price changes conditionally to a given trade imbalance: the response function, directly related to price impact. Within this model, the physical observables measuring such quantities, which we coin as *dressed responses*, are

$$\boldsymbol{R}^{\mathrm{d}} = \mathbb{E}[(\boldsymbol{p} - \boldsymbol{p}_0)(\boldsymbol{y} - \boldsymbol{y}_0)^\top] = \mathbb{C}[\boldsymbol{p}, \boldsymbol{y}]\,, \tag{2.8}$$

$$\boldsymbol{R}_v^{\mathrm{d}} = \mathbb{E}[(\boldsymbol{v} - \boldsymbol{p}_0)(\boldsymbol{y} - \boldsymbol{y}_0)^\top] = \mathbb{C}[\boldsymbol{v}, \boldsymbol{y}]\,. \tag{2.9}$$

Due to the absence of correlations between the uninformed order imbalance $\boldsymbol{u}$ and $\boldsymbol{v}$, one can relate the dressed responses with the responses with respect to $\boldsymbol{x}$, $\boldsymbol{R}$ and $\boldsymbol{R}_v$, which we coin as *bare responses*, through

$$\boldsymbol{R} = \mathbb{E}[(\boldsymbol{p} - \boldsymbol{p}_0)(\boldsymbol{x} - \boldsymbol{y}_0)^\top] = \mathbb{C}[\boldsymbol{p}, \boldsymbol{x}] = \boldsymbol{R}^{\mathrm{d}} - \mathbb{C}[\boldsymbol{p}, \boldsymbol{u}]\,, \tag{2.10}$$

$$\boldsymbol{R}_v = \mathbb{E}[(\boldsymbol{v} - \boldsymbol{p}_0)(\boldsymbol{x} - \boldsymbol{y}_0)^\top] = \mathbb{C}[\boldsymbol{v}, \boldsymbol{x}] = \boldsymbol{R}_v^{\mathrm{d}}\,. \tag{2.11}$$

Interestingly, while the fundamental prices are insensitive to the level of noise trading, the traded prices incorporate some degree of mispricing due to the spurious correlations between the total order imbalance $\boldsymbol{y}$ and the fundamental prices $\boldsymbol{v}$.

3. Equilibrium strategies.

In this section we characterize the equilibrium strategies of the multivariate Kyle model. Although most of the results can be inferred from the seminal work of Caballé and Krishnan [4, 5], we believe it is useful to provide a streamlined version of our own proofs, which we present in a more pedagogical and in some cases more compact form and within which several theoretical issues and empirical implications can be discussed explicitly.

3.1. Linear equilibrium.

Definition 3.1 (linear equilibrium). *By* linear equilibrium*, we mean a set of strategies in which the MM fixes the traded prices* $\boldsymbol{p}$ *and the IT fixes their bid* $\boldsymbol{x}$ *by means of linear rules*

$$\boldsymbol{p} = \boldsymbol{\mu} + \boldsymbol{\Lambda} \boldsymbol{y}\,, \tag{3.1}$$

$$\boldsymbol{x} = \boldsymbol{\alpha} + \boldsymbol{B} \boldsymbol{v}\,, \tag{3.2}$$

such that the two following conditions are satisfied:

1. Profit maximization: *for all alternative strategies with* $\boldsymbol{x}' \neq \boldsymbol{x}$*,*

$$\mathbb{E}[\mathcal{U}_{\mathrm{IT}}(\boldsymbol{x}, \boldsymbol{p})|\boldsymbol{v}] > \mathbb{E}[\mathcal{U}_{\mathrm{IT}}(\boldsymbol{x}', \boldsymbol{p})|\boldsymbol{v}].$$

2. Price efficiency: *The price* $\boldsymbol{p}$ *satisfies* (2.2)*.*

Based on the assumption of a linear strategy for the MM as given in (3.1), we obtain the following results. (Proofs are provided in Appendix A.)

Proposition 3.2. *Imposing the linear pricing rule* (3.1) *for the MM implies that a rational IT will also set the order imbalance* $\boldsymbol{x}$ *as a linear function of the imbalance, with*

$$\boldsymbol{x} = \frac{1}{2}\boldsymbol{\Lambda}_{\mathrm{S}}^{-1}(\boldsymbol{v} - \boldsymbol{\mu})\,, \tag{3.3}$$

where $\boldsymbol{\Lambda}_{\mathrm{S}}$ *denotes the symmetric part of* $\boldsymbol{\Lambda}$*. Furthermore, the profit maximization condition for the IT's strategy implies that* $\boldsymbol{\Lambda}_{\mathrm{S}}$ *has to be positive-definite (PD).*

Equation (3.1) shows that $\boldsymbol{\Lambda}$ plays the role of adjusting the traded price level $\boldsymbol{p}$ proportionally to the imbalance $\boldsymbol{y}$. Proposition 3.2 shows that $\boldsymbol{\Lambda}$ also plays the role of setting the order imbalance from the informed trader $\boldsymbol{x}$ given the knowledge of the fundamental price $\boldsymbol{v}$. Moreover, (3.3) together with (3.1) implies that both the order imbalance $\boldsymbol{y}$ and the traded price $\boldsymbol{p}$ are normally distributed random variables, due to stability of the Gaussian distribution under convolution. This last property is at the core of the following proposition.

Proposition 3.3. *Assuming a linear strategy for the IT as in* (3.3) *implies that the parameters of the pricing rule of* (3.1) *for the MM are given by*

$$\begin{aligned} \boldsymbol{\mu} &= \boldsymbol{p}_0 - \boldsymbol{\Lambda}\boldsymbol{y}_0\,, \\ \boldsymbol{\Lambda} &= \boldsymbol{R}_v^{\mathrm{d}}(\boldsymbol{\Omega}^{\mathrm{d}})^{-1}. \end{aligned} \tag{3.4}$$

Note that $\boldsymbol{y}_0$, $\boldsymbol{R}_v^{\mathrm{d}}$, and $\boldsymbol{\Omega}^{\mathrm{d}}$ depend on $\boldsymbol{x}$, and therefore on $\boldsymbol{\Lambda}_{\mathrm{S}}$. Substituting (3.4) into (3.3) allows us to close the system and find an equation for $\boldsymbol{\Lambda}$. Note, however, that the as obtained system has many possible solutions. To constrain the latter we must impose the profit maximization condition introduced in Proposition 3.2.

Proposition 3.4. *Assume that there exists a solution to the utility maximization problem of the IT of the form given by* (3.3) *and to the pricing rule for the MM given by* (3.1) *and* (3.4)*. Then, the profit maximization condition that* $\boldsymbol{\Lambda}_{\mathrm{S}}$ *has to be PD implies that* $\boldsymbol{\Lambda}$ *is symmetric and satisfies the equation*

$$\frac{1}{4}\boldsymbol{\Sigma}_0 = \boldsymbol{\Lambda}\boldsymbol{\Omega}\boldsymbol{\Lambda}\,. \tag{3.5}$$

The symmetry of $\boldsymbol{\Lambda}$ leads to a unique solution that can be expressed in terms of the parameters of the problem (i.e., $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Omega}$). Denoting by $\sqrt{\boldsymbol{Y}}$ the unique (see Lemma A.1 in Appendix A.3) PD solution of the matrix equation $\boldsymbol{X}\boldsymbol{X} = \boldsymbol{Y}$ for a SPD $\boldsymbol{Y}$ we introduce the following theorem.

Theorem 3.5 (existence and unicity of the linear equilibrium). *There exists a unique linear equilibrium given by the strategies* (3.1) *and* (3.3)*, where*

$$\boldsymbol{\Lambda} = \tfrac{1}{2}\mathcal{R}^{-1}\sqrt{\mathcal{R}\boldsymbol{\Sigma}_0\mathcal{L}}\mathcal{L}^{-1}\,, \tag{3.6}$$

$$\boldsymbol{y}_0 = 0\,, \tag{3.7}$$

$$\boldsymbol{\mu} = \boldsymbol{p}_0\,. \tag{3.8}$$

Here $\mathcal{L}, \mathcal{R}$ *are a factorization of the matrix* $\boldsymbol{\Omega} = \mathcal{L}\mathcal{R}$ *satisfying* $\mathcal{L} = \mathcal{R}^{\top}$*.*

Note that the existence result is a special case of [4, Proposition 3.1], specialized to the case of a single IT with perfect information, and the absence of other equilibria with nonsymmetric $\mathbf{\Lambda}$ is proven in the unpublished preprint [5]. Because the symmetry property of $\mathbf{\Lambda}$ has crucial consequences on both price formation and pricing (see the discussion in sections 4 and 5), ruling out the existence of nonsymmetric equilibria is an important step in the analysis of the multivariate Kyle model.

Example 1 (solution of the univariate Kyle model). The specialization of Theorem 3.5 to the $n = 1$ case in which only one asset is traded yields the well-known solution derived in [9] (light lowercase symbols are scalar versions of the bold uppercase and bold lowercase ones),

$$\lambda = \frac{1}{2}\sqrt{\frac{\sigma_0}{\omega}}\,, \tag{3.9}$$

indicating that the constant of proportionality between price and imbalances scales with the amount of price fluctuations $\sqrt{\sigma_0}$ and is inversely proportional to the typical fluctuations of the noise $\sqrt{\omega}$. Intuitively, the larger price deviations the rational MM expects, the more weight he should give to volume imbalances in order to forecast the fundamental price to the imbalances. On the other hand, the more noise there is in the system, the less the volume signal is reliable, so that the traded price is closer to the uninformed prior p_0.

It is important to note that (3.5) alone does not imply the symmetry of $\mathbf{\Lambda}$, as it is necessary to further impose PDness in order to obtain symmetry. In fact, there are symmetric solutions to (3.5) that are not PD. These solutions lead to efficient traded prices but they do not optimize the IT's utility.

Example 2 (saddle point solutions). Let $\boldsymbol{x}$ be a strategy of the form given in (3.3) and let $\boldsymbol{\mu} = \boldsymbol{p}_0$ and let $\mathbf{\Lambda}$ be a symmetric but not PD solution of (3.5). By construction, the gradient of $\mathbb{E}[\mathcal{U}_{IT}]$ with respect to $\boldsymbol{x}$ at this point is zero, satisfying the first order condition for the IT, but this does not guarantee that it is a maximum. We now take an arbitrary perturbation of the strategy $\boldsymbol{x} + \delta\boldsymbol{x}$ and compute the difference in utility between the two strategies:

$$\begin{aligned}
\delta\mathbb{E}[\mathcal{U}_{IT}] &= \mathbb{E}[\mathcal{U}_{IT}](\boldsymbol{x} + \delta\boldsymbol{x}) - \mathbb{E}[\mathcal{U}_{IT}](\boldsymbol{x}) \\
&= \mathbb{E}[\delta\boldsymbol{x}^\top(\boldsymbol{v} - \boldsymbol{p}_0) - 2\boldsymbol{x}^\top\mathbf{\Lambda}\delta\boldsymbol{x} - \delta\boldsymbol{x}^\top\mathbf{\Lambda}\delta\boldsymbol{x}] \\
&= -\delta\boldsymbol{x}^\top\mathbf{\Lambda}\delta\boldsymbol{x}\,.
\end{aligned}$$

One can see that $\delta\mathbb{E}[\mathcal{U}_{IT}]$ is always negative for arbitrary $\delta\boldsymbol{x}$ if and only if $\mathbf{\Lambda}$ is PD. Without the second order condition the IT could find better strategies than $\boldsymbol{x}$, meaning that the solution that corresponds to that $\mathbf{\Lambda}$ is not an equilibrium but a saddle point. In fact, when $\mathbf{\Lambda}$ is not PD, $\mathbb{E}[\mathcal{U}_{IT}]$ can be arbitrarily large.

In such a case, the IT would trade arbitrarily large quantities of some linear combinations of assets, and no equilibrium would be possible. Imposing a PD impact matrix $\mathbf{\Lambda}$ allows the MM to fix the optimal strategy of the IT, which otherwise would be undetermined—at least in the absence of a further risk penalty in the IT utility function.

3.2. Price and volume covariances of the equilibrium solution. We now investigate what are the testable implications of the linear equilibrium described above. Here, we provide the relations between the observables that one would measure in such a situation.

Proposition 3.6 (inconspicous equilibrium). *In the equilibrium, the expectation of the dressed order imbalance is equal to*

$$\boldsymbol{y}_0 = \mathbb{E}[\boldsymbol{y}] = \mathbb{E}[\boldsymbol{u}] = 0\,, \tag{3.10}$$

implying that the IT hides her trades among the noise traders ($\mathbb{E}[\boldsymbol{x}] = 0$)*. The IT order imbalance covariance can be expressed as*

$$\mathbb{C}[\boldsymbol{x},\boldsymbol{x}] = \boldsymbol{\Omega} \tag{3.11}$$

and

$$\boldsymbol{\Omega}^{\mathrm{d}} = 2\boldsymbol{\Omega}\,. \tag{3.12}$$

Proof. The results in (3.10) and (3.11) are obtained by directly evaluating the expectations:

$$\begin{aligned}\mathbb{E}[\boldsymbol{x}] &= \frac{1}{2}\boldsymbol{\Lambda}^{-1}\mathbb{E}[\boldsymbol{v}-\boldsymbol{p}_0] = 0\,,\\ \mathbb{C}[\boldsymbol{x},\boldsymbol{x}] &= \frac{1}{4}\boldsymbol{\Lambda}^{-1}\mathbb{E}[(\boldsymbol{v}-\boldsymbol{p}_0)(\boldsymbol{v}-\boldsymbol{p}_0)^\top]\boldsymbol{\Lambda}^{-1} = \frac{1}{4}\boldsymbol{\Lambda}^{-1}\boldsymbol{\Sigma}_0\boldsymbol{\Lambda}^{-1} = \boldsymbol{\Omega}\,.\end{aligned}$$

The last result follows trivially from (3.11) and the definition of $\boldsymbol{\Omega}^{\mathrm{d}}$. ■

This result implies that, in addition to matching the trading level of the NT in order to conceal his information (as obtained in the one-dimensional case), the IT is required to match the *directions* in which the NT trades (i.e., the eigenvectors of $\boldsymbol{\Omega}$) such that the trading behavior of the IT is indistinguishable from that of the NT. An even more interesting consequence of this behavior is given in the next proposition.

Proposition 3.7 (information diffusion). *The fundamental price covariance and the traded price covariance in the equilibrium are related by*

$$\boldsymbol{\Sigma} = \frac{1}{2}\boldsymbol{\Sigma}_0\,. \tag{3.13}$$

Furthermore, the residual information is

$$\begin{aligned}\mathbb{C}[\boldsymbol{v},\boldsymbol{v}|\boldsymbol{p}] &= \mathbb{C}[\boldsymbol{v},\boldsymbol{v}] - \mathbb{C}[\boldsymbol{v},\boldsymbol{p}]\mathbb{C}[\boldsymbol{p},\boldsymbol{p}]^{-1}\mathbb{C}[\boldsymbol{p},\boldsymbol{v}]\\ &= \boldsymbol{\Sigma}_0 - \boldsymbol{B}\boldsymbol{\Lambda}\boldsymbol{\Sigma}_0\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_0\boldsymbol{\Lambda}^\top\boldsymbol{B}^\top,\end{aligned} \tag{3.14}$$

which at equilibrium reads

$$\mathbb{C}[\boldsymbol{v},\boldsymbol{v}|\boldsymbol{p}] = \frac{1}{2}\boldsymbol{\Sigma}_0. \tag{3.15}$$

Proof. The traded price covariance is obtained as

$$\boldsymbol{\Sigma} = \mathbb{C}[\boldsymbol{p},\boldsymbol{p}] = \boldsymbol{\Lambda}\boldsymbol{\Omega}^{\mathrm{d}}\boldsymbol{\Lambda}^\top = 2\boldsymbol{\Lambda}\boldsymbol{\Omega}\boldsymbol{\Lambda} = \frac{1}{2}\boldsymbol{\Sigma}_0\,, \tag{3.16}$$

where the last equality comes from (A.14) in Appendix A.3. To find the residual information one needs to compute the covariance between traded prices and fundamental prices,

$$\mathbb{C}[\boldsymbol{v},\boldsymbol{p}] = \boldsymbol{B}\boldsymbol{\Lambda}\mathbb{C}[\boldsymbol{v},\boldsymbol{v}] = \boldsymbol{B}\boldsymbol{\Lambda}\boldsymbol{\Sigma}_0,$$

and later use the fact that at equilibrium $\boldsymbol{B}\boldsymbol{\Lambda} = \frac{1}{2}$. ∎

While the fact that the two covariances are related is expected, the fact that they are proportional to one another is nontrivial. Furthermore (3.16) implies that the role of the impact matrix $\boldsymbol{\Lambda}$ is to "rotate" the fluctuations of the volume in the direction of the fluctuations of the fundamental price. The intuition behind this behavior is that due to camouflage the IT will trade in the same direction as the NT. Hence, a rational MM trying to enforce price efficiency will be required to convert fluctuations along the principal components of $\boldsymbol{\Omega}$ into forecasts of price fluctuations along the principal components of $\boldsymbol{\Sigma}_0$, thus matching the expected covariance $\boldsymbol{\Sigma}_0$. The fact that the proportionality constant is equal to $1/2$ (only half of the information is revealed) is of minor importance and is a consequence of the single-period feature of the model. Extending the Kyle framework to multiple time steps would turn the constant factor $1/2$ into a time-varying function expressing the rate at which information is incorporated into prices. (See the original Kyle paper [9] for the single asset case and [15, 10] for, respectively, the discrete and continuous multiasset case.)

Finally, it is interesting to characterize the responses in the context of a multivariate Kyle model.

Proposition 3.8 (responses). *The propositions above and* (2.10) *imply that*

$$\boldsymbol{R}_v = \boldsymbol{R}_v^{\mathrm{d}} = \boldsymbol{R}^{\mathrm{d}} = 2\boldsymbol{R} = \frac{1}{2}\boldsymbol{\Sigma}_0\boldsymbol{\Lambda}^{-1}\,, \tag{3.17}$$

so that the responses are uniquely determined by $\boldsymbol{\Sigma}_0$ *and* $\boldsymbol{\Lambda}$.

Proof. The equilibrium values of the responses are found by plugging the equilibrium strategies into their definitions:

$$\begin{aligned}
\boldsymbol{R}_v^{\mathrm{d}} &= \boldsymbol{R}_v = \mathbb{E}[(\boldsymbol{v}-\boldsymbol{p}_0)\boldsymbol{x}^\top] = \frac{1}{2}\mathbb{E}[(\boldsymbol{v}-\boldsymbol{p}_0)(\boldsymbol{v}-\boldsymbol{p}_0)^\top]\boldsymbol{\Lambda}^{-1} = \frac{1}{2}\boldsymbol{\Sigma}_0\boldsymbol{\Lambda}^{-1}\,,\\
\boldsymbol{R}^{\mathrm{d}} &= \mathbb{E}[(\boldsymbol{p}-\boldsymbol{p}_0)\boldsymbol{y}^\top] = \boldsymbol{\Lambda}\mathbb{E}[\boldsymbol{y}\boldsymbol{y}^\top] = 2\boldsymbol{\Lambda}\boldsymbol{\Omega} = \frac{1}{2}\boldsymbol{\Sigma}_0\boldsymbol{\Lambda}^{-1}\,,\\
\boldsymbol{R} &= \mathbb{E}[(\boldsymbol{p}-\boldsymbol{p}_0)\boldsymbol{x}^\top] = \boldsymbol{\Lambda}\mathbb{E}[\boldsymbol{x}\boldsymbol{x}^\top] = \boldsymbol{\Lambda}\boldsymbol{\Omega} = \frac{1}{4}\boldsymbol{\Sigma}_0\boldsymbol{\Lambda}^{-1}\,.
\end{aligned}$$
∎

In section 5 we will see how this naturally leads to the calibration of an impact model in which part of the estimation noise can be reduced by imposing the response structure above as a reasonable prior. Finally, note that while the dressed and bare responses of the fundamental price coincide, the dressed response of the traded price is a factor 2 larger than the bare response, due to the spurious correlations between prices and volume transiently introduced by the NT.

3.3. Utilities and competitive market making. In this model the utilities of the three agents are such that they sum to zero:

$$\begin{aligned}\mathcal{U}_{IT} &= -\boldsymbol{x}^\top(\boldsymbol{p}-\boldsymbol{v})\,,\\ \mathcal{U}_{NT} &= -\boldsymbol{u}^\top(\boldsymbol{p}-\boldsymbol{v})\,,\\ \mathcal{U}_{MM} &= (\boldsymbol{x}+\boldsymbol{u})^\top(\boldsymbol{p}-\boldsymbol{v})\,.\end{aligned}$$

It is thus interesting to characterize how utilities are transferred from one agent to the other at equilibrium.

Proposition 3.9 (utilities at equilibrium). *The utilities of the three agents at equilibrium are equal to*

$$\mathbb{E}[\mathcal{U}_{IT}] = \operatorname{tr}(\boldsymbol{R})\,, \tag{3.18}$$

$$\mathbb{E}[\mathcal{U}_{NT}] = -\operatorname{tr}(\boldsymbol{R})\,, \tag{3.19}$$

$$\mathbb{E}[\mathcal{U}_{MM}] = 0\,. \tag{3.20}$$

Proof.

$$\begin{aligned}\mathbb{E}[\mathcal{U}_{IT}] &= -\mathbb{E}[\boldsymbol{x}^\top(\boldsymbol{p}-\boldsymbol{v})] = \operatorname{tr}(\mathbb{E}[\boldsymbol{v}\boldsymbol{x}^\top] - \mathbb{E}[\boldsymbol{p}\boldsymbol{x}^\top]) = \operatorname{tr}(\boldsymbol{R}_v - \boldsymbol{R}) = \operatorname{tr}(\boldsymbol{R})\,,\\ \mathbb{E}[\mathcal{U}_{NT}] &= -\mathbb{E}[\boldsymbol{u}^\top(\boldsymbol{p}-\boldsymbol{v})] = \operatorname{tr}(\mathbb{E}[\boldsymbol{v}\boldsymbol{u}^\top] - \mathbb{E}[\boldsymbol{p}\boldsymbol{u}^\top]) = \operatorname{tr}(0 - \boldsymbol{R}) = -\operatorname{tr}(\boldsymbol{R})\,,\\ \mathbb{E}[\mathcal{U}_{MM}] &= -\mathbb{E}[\mathcal{U}_{IT}] - \mathbb{E}[\mathcal{U}_{NT}] = 0\,.\end{aligned}$$

■

As in the standard Kyle model, the linear equilibrium of the multivariate Kyle model is such that wealth is transferred from the NT to the IT, who is able to capitalize on his informational advantage. The role of the MM is more subtle: we assumed that the MM enforces an efficient price, and hence by construction he is not optimizing his wealth. However, as we show in Proposition 3.9, imposing price efficiency through the pricing rule (2.2) results in a break-even for the MM in the sense that the expectation of his utility is 0. In the next example we show that the converse is not true: imposing an unconditional break-even condition for the MM is in general not sufficient to ensure efficient prices, and one needs to impose a conditional, assetwise condition in order to recover the pricing rule (2.2).

Example 3 (efficient prices and unconditional break-even). Consider the unconditional break-even condition $\mathbb{E}[\mathcal{U}_{MM}] = 0$. The expected utility of the MM in the univariate model reads

$$\mathbb{E}[\mathcal{U}_{MM}] = \frac{1}{2}(\mu - p_0)^2\lambda^{-1} + \lambda\omega^{\mathrm{d}} - r_v^{\mathrm{d}} = 0.$$

Imposing $\mu = p_0$ rules out the presence of complex solutions regardless of r_v^{d} and ω^{d} and therefore

$$\begin{aligned}\lambda &= r_v^d/\omega^d\,,\\ \mu &= p_0\,.\end{aligned}$$

Hence in one dimension a break-even condition for the MM, together with the requirement $\mu = p_0$, is equivalent to imposing price efficiency. This equivalence stems from the low dimensionality of the system because when we impose the break-even condition in one dimension we obtain one equation for one parameter. (μ is fixed by imposing that the solution has to be real for any choice of parameters.)

In the multivariate Kyle model this is not the case. In particular, the unconditional break-even condition writes

$$\mathbb{E}[\mathcal{U}_{MM}] = \operatorname{tr}\left(\frac{1}{2}(\boldsymbol{\mu} - \boldsymbol{p}_0)(\boldsymbol{\mu} - \boldsymbol{p}_0)^\top \boldsymbol{\Lambda}_{\mathrm{S}}^{-1} + \boldsymbol{\Lambda}\boldsymbol{\Omega}^{\mathrm{d}} - \boldsymbol{R}_v\right) = 0\,.$$

In this case we obtain again one equation but we now need to fix $\sim n^2$ parameters. This means that there are many ways to break-even while only one is efficient. In a more restrictive setting where the MM has to break-even for each asset individually, the condition can be written as

$$\operatorname{diag}\left(\frac{1}{2}(\boldsymbol{\mu} - \boldsymbol{p}_0)(\boldsymbol{\mu} - \boldsymbol{p}_0)^\top \boldsymbol{\Lambda}_{\mathrm{S}}^{-1} + \boldsymbol{\Lambda}\boldsymbol{\Omega}^{\mathrm{d}} - \boldsymbol{R}_v\right) = 0. \tag{3.21}$$

In this case there are n equations, which are still not enough to fix $\sim n^2$ parameters.

Therefore, in one dimension the condition $\mathbb{E}[\mathcal{U}_{MM}] = 0$ is both a necessary and sufficient condition for the prices to be efficient. In multiple dimensions, although $\mathbb{E}[\mathcal{U}_{MM}] = 0$ is still a consequence of price efficiency, the converse is not true, as clarified in the next remark.

Remark 1 (sufficient and necessary conditions for efficiency). The assetwise unconditional break-even (3.21) can be also written as

$$\mathbb{E}[(p_i - v_i)y_i] = 0 \tag{3.22}$$

for each asset i. Under the assumption $\boldsymbol{\mu} = \boldsymbol{p}_0$ and in matrix form, (3.22) leads to

$$\begin{aligned} 0 &= \operatorname{diag}(\mathbb{E}[(\boldsymbol{p} - \boldsymbol{v})\boldsymbol{y}^\top]) \\ &= \operatorname{diag}(\boldsymbol{\Lambda}\mathbb{E}[\boldsymbol{y}\boldsymbol{y}^\top] - \mathbb{E}[\boldsymbol{v}\boldsymbol{y}^\top]) \\ &= \operatorname{diag}(\boldsymbol{\Lambda}\boldsymbol{\Omega}^{\mathrm{d}} - \boldsymbol{R}_v), \end{aligned}$$

which corresponds to the diagonal of (3.4). Yet, this condition is not strong enough to recover the pricing rule (2.2) as it allows us to fix n parameters, whereas we would need to fix n^2 of them.

If instead we consider a conditional break-even (that would naturally arise from a Bertrand type of auction), we recover as expected the martingale condition (2.2):

$$\mathbb{E}[(p_i - v_i)y_i|\boldsymbol{y}] = 0, \qquad \text{which implies} \qquad y_i\mathbb{E}[(p_i - v_i)|\boldsymbol{y}] = 0.$$

Hence, unlike in one dimension (where it is not necessary to assume that the break-even condition is enforced conditionally), in n dimensions it is required that the MM

(i) breaks even on every asset and
(ii) breaks even *regardless* of the particular realization of the order imbalance $\boldsymbol{y}$.

This can be rationalized by noting that an unconditional break-even condition would imply that for some realizations of $\boldsymbol{y}$ the MM would consistently lose with respect to some assets and win with respect to others and this would provide an incentive to modify the strategy on the losing bets. Therefore strategies in which the break-even condition is (i) conditionally violated or (ii) does not hold on each asset would not be equilibria.

4. Interpretation. The properties of the linear equilibrium that have been illustrated in the section above provide some intuition about the phenomenology of the model, but they do not help in making sense of (3.6) for the impact $\boldsymbol{\Lambda}$. The following discussions are meant to provide a justification of the expression for $\boldsymbol{\Lambda}$, so as to shed some light on its structure. As far as we know, the following arguments have never appeared in the literature before.

4.1. Whitening of prices and volumes. The derivation of the linear equilibrium presented in Theorem 3.5 relies on the quadratic matrix equation (3.5), which is solved by (see the proof of Proposition A.2)

$$\boldsymbol{\Lambda} = \frac{1}{2}\mathcal{G}\boldsymbol{O}\mathcal{L}^{-1}\,, \tag{4.1}$$

where we have used a decomposition for the price correlations $\boldsymbol{\Sigma}_0 = \mathcal{G}\mathcal{D}$ analogous to that of $\boldsymbol{\Omega} = \mathcal{L}\mathcal{R}$ and $\boldsymbol{O}$ is an orthogonal matrix. The remarkable part of this finding is that, regardless of $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Omega}$, there always exists a unique change of basis $\boldsymbol{O}$ such that $\boldsymbol{\Lambda}$ is SPD. Hence, the prediction process operated by the MM that results in the pricing rule $\Delta\boldsymbol{v} = \boldsymbol{v} - \boldsymbol{p}_0 = \boldsymbol{\Lambda}\boldsymbol{y}$ can be seen as resulting from the following steps:

1. Apply the *whitening* transformation $\mathcal{L}^{-1}$ to $\boldsymbol{y}$, so as to obtain the whitened imbalance $\tilde{\boldsymbol{y}} = 2^{-1/2}\mathcal{L}^{-1}\boldsymbol{y}$ with $\mathbb{C}[\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{y}}] = \mathbb{I}$.
2. Apply the rotation $\boldsymbol{O}$ to $\tilde{\boldsymbol{y}}$, obtaining the whitened fundamental price variations $\Delta\tilde{\boldsymbol{v}} = \boldsymbol{O}\tilde{\boldsymbol{y}}$. As before, one has in fact $\mathbb{C}[\Delta\tilde{\boldsymbol{v}}, \Delta\tilde{\boldsymbol{v}}] = \mathbb{E}[\Delta\tilde{\boldsymbol{v}}\Delta\tilde{\boldsymbol{v}}^\top] = \mathbb{I}$.
3. Apply the *inverse whitening* transformation $\mathcal{G}$ in order to find the prediction for the fundamental price variation $\Delta\boldsymbol{v} = \boldsymbol{v} - \boldsymbol{p}_0 = 2^{-1/2}\mathcal{G}\Delta\tilde{\boldsymbol{v}}$.

While the first and the third steps of this procedure are intuitive and can be seen as arising from dimensional analysis only, the nature of the rotation $\boldsymbol{O}$ applied during the second step is less trivial and will be investigated more closely in the following sections.

Example 4 (degeneracy in the one-dimensional Kyle model). In one dimension (3.5) becomes

$$\frac{\sigma_0}{4} = \lambda^2\omega\,,$$

leading to the couple of solutions $\lambda = \pm\frac{1}{2}\sqrt{\sigma_0/\omega}$. The degeneracy between the solutions is easily resolved by imposing positive-definiteness of the impact $\boldsymbol{\Lambda}$, which selects the positive solution. In this case we have $\mathcal{G} = \sqrt{\sigma_0}$, $\mathcal{L} = \sqrt{\omega}$, $\boldsymbol{O} = 1$. Hence, the one-dimensional Kyle model has too low a dimensionality for anything nontrivial to happen from the point of view of $\boldsymbol{O}$: it is only in a higher dimension that one can appreciate its general structure.

4.2. Basis of prices and basis of volumes. As we show in the proof of Proposition A.2, the rotation $\boldsymbol{O}$ that appears in (4.1) can be expressed as

$$\boldsymbol{O} = (\mathcal{R}\mathcal{G})^{-1}\sqrt{(\mathcal{R}\mathcal{G})(\mathcal{R}\mathcal{G})^\top}\,. \tag{4.2}$$

In order to make some progress in understanding the complex nature of this rotation for $n > 1$, it is important to notice that it is completely specified in terms of the matrix $\mathcal{R}\mathcal{G}$. Let us assume that the matrix factorization chosen to compute $\boldsymbol{\Omega}$ and $\boldsymbol{\Sigma}_0$ is the principal component decomposition:

$$\boldsymbol{\Omega} = \underbrace{\boldsymbol{W}\operatorname{diag}(\sqrt{\boldsymbol{\omega}})}_{\mathcal{L}}\underbrace{\operatorname{diag}(\sqrt{\boldsymbol{\omega}})\boldsymbol{W}^\top}_{\mathcal{R}}, \tag{4.3}$$

$$\boldsymbol{\Sigma}_0 = \underbrace{\boldsymbol{S}\operatorname{diag}(\sqrt{\boldsymbol{\sigma}_0})}_{\mathcal{G}}\underbrace{\operatorname{diag}(\sqrt{\boldsymbol{\sigma}_0})\boldsymbol{S}^\top}_{\mathcal{D}}, \tag{4.4}$$

where $\boldsymbol{W}$ and $\boldsymbol{S}$ are orthogonal matrices, and where the vectors $\sqrt{\boldsymbol{\omega}}$ and $\sqrt{\boldsymbol{\sigma}_0}$ have positive elements. Then we have

$$\mathcal{R}\mathcal{G} = \operatorname{diag}(\sqrt{\boldsymbol{\omega}})\boldsymbol{W}^\top\boldsymbol{S}\operatorname{diag}(\sqrt{\boldsymbol{\sigma}_0}), \tag{4.5}$$

indicating that, besides the diagonal matrices $\operatorname{diag}(\sqrt{\boldsymbol{\omega}})$ and $\operatorname{diag}(\sqrt{\boldsymbol{\sigma}_0})$ that set the scale of the fluctuations of fundamental prices and order imbalancees, it is the overlap $\boldsymbol{W}^\top\boldsymbol{S}$ that links the eigenvectors of the volumes with that of the fundamental price. In order to obtain a stronger insight about the structure of $\boldsymbol{O}$, we shall proceed with some simple examples.

Example 5 (two-dimensional Kyle model). In two dimensions the matrix $\boldsymbol{O}$ can be characterized, without loss of generality, by a single angle θ:[2]

$$\boldsymbol{O} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Consider a system given by the following correlations:

$$\boldsymbol{\Sigma}_0 = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \qquad \boldsymbol{\Omega} = \begin{pmatrix} \omega_1 & 0 \\ 0 & \omega_2 \end{pmatrix}.$$

In this case we can calculate $\boldsymbol{O}$ explicitly. In particular one has

$$\mathcal{G} = \frac{1}{\sqrt{2}}\begin{pmatrix} \sqrt{1+\rho} & -\sqrt{1-\rho} \\ \sqrt{1+\rho} & \sqrt{1-\rho} \end{pmatrix}, \qquad \mathcal{L}^{-1} = \begin{pmatrix} \omega_1^{-1/2} & 0 \\ 0 & \omega_2^{-1/2} \end{pmatrix},$$

and imposing the symmetry of $\boldsymbol{\Lambda}$ one can show that

$$\theta = \arcsin\left(\frac{\omega_1^{-1/2}\sqrt{1+\rho} + \omega_2^{-1/2}\sqrt{1-\rho}}{\sqrt{2}\Delta}\right),$$

where $\Delta = \sqrt{\omega_1 + \omega_2 + 2(\omega_1\omega_2)^{1/2}\sqrt{1-\rho^2}}$. Finally we obtain

$$\boldsymbol{\Lambda} = \frac{1}{2\Delta}\begin{pmatrix} 1 + \sqrt{\frac{\omega_2}{\omega_1}}\sqrt{1-\rho^2} & \rho \\ \rho & 1 + \sqrt{\frac{\omega_1}{\omega_2}}\sqrt{1-\rho^2} \end{pmatrix}. \tag{4.6}$$

[2] Indeed, by consistently choosing the factorization of $\mathcal{G}$ and $\mathcal{L}$ one can fix arbitrarily the determinant of $\boldsymbol{O}$ to be plus or minus one due to (4.2).

We can use the results in Example 5 to examine a couple of extreme cases: the one of extreme illiquidity and the one in which assets are strong correlations.

Example 6 (extreme illiquidity). Let us consider the limit in which $\omega_2 = \epsilon\omega$ with $\epsilon \to 0$, whereas $\omega_1 = \omega$, implying that $y_2\epsilon^{-1/2}$ is expected to be finite. Then the prediction of the MM up to first order in ϵ is

$$\boldsymbol{p} - \boldsymbol{p}_0 = \frac{1}{2\sqrt{\omega}} \begin{pmatrix} y_1 \\ \rho y_1 + \sqrt{1-\rho^2}(y_2\epsilon^{-1/2}) \end{pmatrix},$$

implying that the efficient price of a liquid instrument is fixed solely by the volume traded on that same instrument, whereas for illiquid markets it is important to take into account quantities traded on liquid, correlated markets. On the other hand, the behavior of the IT up to first order in ϵ is given by

$$\boldsymbol{x} = \sqrt{\omega} \begin{pmatrix} \Delta v_1 \\ \sqrt{\frac{\epsilon}{1-\rho^2}}(\Delta v_2 - \rho \Delta v_1) \end{pmatrix}.$$

Now the IT is encouraged to trade only the liquid asset and furthermore ignore the illiquid one when placing the bid. Similarly to the previous case, the traded price will remain of order one because the market orders for the second asset will be order $\sqrt{\epsilon}$.

Price innovation in Example 6 can be interpreted as an indication of how an efficient market should work: the symmetry of cross-impact implies that the effect of trading one dollar of a very liquid asset a_1 on the price of an iliquid asset a_2 is the same as the effect of trading a dollar of asset a_2 on the price of asset a_1, but because there are going to be very few dollars traded on a_2 it will not affect significantly the price of a_1. Conversely, the price of a_2 will be heavily driven by the traded volume of a_1.

Although we only proved the existence of the linear equilibrium for invertible $\boldsymbol{\Sigma}_0$, we can study the behavior of $\boldsymbol{\Lambda}$ in the limiting case in which $\boldsymbol{\Sigma}_0$ has low rank as in the two following examples.

Example 7 (strongly correlated prices). Consider $\rho = 1 - \delta$ with $\delta \to 0$, so that $(\Delta p_1 - \Delta p_2)/\sqrt{\delta}$ is finite. According to (4.6) the prediction of the MM up to first order in δ is

$$\boldsymbol{p} - \boldsymbol{p}_0 = \frac{1}{2\sqrt{\omega_1 + \omega_2}} \begin{pmatrix} y_1 + y_2 \\ y_1 + y_2 \end{pmatrix},$$

implying that when dealing with strongly correlated instruments, the efficient prices can be built by summing algebraically the volumes traded on each of them before normalizing by the global liquidity. On the other hand, the bet of the IT up to first order in δ can be written as

$$\boldsymbol{x} = \frac{1}{2}\sqrt{\frac{1}{\omega_1 + \omega_2}} \begin{pmatrix} \omega_1(\Delta v_1 + \Delta v_2) \\ \omega_2(\Delta v_1 + \Delta v_2) \end{pmatrix} + \sqrt{\frac{\omega_1\omega_2}{2(\omega_1 + \omega_2)}} \begin{pmatrix} \frac{1}{\sqrt{\delta}}(\Delta v_1 - \Delta v_2) \\ \frac{1}{\sqrt{\delta}}(\Delta v_2 - \Delta v_1) \end{pmatrix}.$$

In this case the IT is encouraged to place orders proportional to either the average price variation, or orders of finite size proportional to the relative difference, rescaled by the typical size of the relative price moves $\sqrt{\delta}$. Indeed, the IT should bet on price variations inversely proportional to how common they are.

The specialization of the result above to the extreme case $\rho \to \pm 1$ regime is particularly interesting, because it is a consequence of a more general result that can be applied whenever $\boldsymbol{\Sigma}_0$ is rank one, which is shown in the next example.

Example 8 (rank one price covariance). Let us consider the case in which all eigenvalues of $\boldsymbol{\Sigma}_0$ except for one tend to 0, and therefore $\boldsymbol{\Sigma}_0$ tends to a rank 1 matrix. The interest of this example resides in the fact that $\boldsymbol{\Lambda}$ displays in this case a particularly illuminating expression, providing a simple recipe for pooling together volumes belonging to different financial instruments correlated to the same underlying product. Without loss of generality, $\boldsymbol{\Sigma}_0$ in the rank one case can be written as $\boldsymbol{\Sigma}_0 = \boldsymbol{s}\sigma\boldsymbol{s}^\top$, where $\boldsymbol{s} \in \mathbb{R}^n$ is a unit vector and $\sigma > 0$. It is then straightforward to verify that the matrix

$$\boldsymbol{\Lambda} = \frac{1}{2}\boldsymbol{s}\left(\frac{\sigma}{\boldsymbol{s}^\top\boldsymbol{\Omega}\boldsymbol{s}}\right)^{1/2}\boldsymbol{s}^\top \tag{4.7}$$

yields the linear equilibrium of the model. The above equation has a very simple interpretation:

- The matrix $\boldsymbol{\Lambda}$ commutes with $\boldsymbol{\Sigma}_0$, and its eigenvectors are insensitive to those of $\boldsymbol{\Omega}$ (i.e., the volume fluctuations induced by noise traders).
- The factor $\sigma^{1/2}$ sets the scale for the price variations as being that of the only mode of $\boldsymbol{\Sigma}_0$ that has nonzero fluctuations.
- By writing the principal component decomposition of $\boldsymbol{\Omega}$ as $\boldsymbol{\Omega} = \sum_{a=1}^n \boldsymbol{w}_a\omega_a\boldsymbol{w}_a^\top$, one can write the denominator as

$$(\boldsymbol{s}^\top\boldsymbol{\Omega}\boldsymbol{s})^{1/2} = \left(\sum_{a=1}^n(\boldsymbol{w}_a^\top\boldsymbol{s})^2\omega_a\right)^{1/2}. \tag{4.8}$$

  Such a denominator sets the scaling with volume of the impact $\boldsymbol{\Lambda}$. It amounts to a projection of $\boldsymbol{\Omega}$ on the only nonzero mode of $\boldsymbol{\Sigma}_0$, or equivalently a sum of the individual volume variances ω_a of $\boldsymbol{\Omega}$, weighted by their projections on $\boldsymbol{s}$.

5. Implications. The expression (3.6) for the impact matrix $\boldsymbol{\Lambda}$ is of interest beyond the context of the Kyle model. It provides an insightful inference prescription for cross-impact models, as opposed to what has been done so far in the context of cross-impact fitting (see [2, 12, 16, 14]). In such a context, one is interested in estimating from empirical data a model to forecast a price variation $\Delta\boldsymbol{p}$ with a predictor of the form

$$\hat{\Delta}\boldsymbol{p} = \hat{\boldsymbol{\Lambda}}\boldsymbol{y}\,, \tag{5.1}$$

in which it is essential to faithfully model how trading the instrument i will impact the instrument j. The goal of this section is to establish a link between the Kyle model and the empirical calibration of (5.1), where the price variations $\Delta\boldsymbol{p}$ and the imbalances $\boldsymbol{y}$ are sampled from empirical data.

5.1. From theory to data: Empirical averages and loss function. First, let us consider a dataset of T independent and identically distributed efficient price variations and volumes, $\{\Delta\boldsymbol{p}^{(t)}\}_{t=1}^T$ and $\{\boldsymbol{y}^{(t)}\}_{t=1}^T$, sampled from an unknown distribution with bounded variance that needs not be related to the Kyle model presented in the previous sections of the manuscript.

Accordingly, with a slight abuse of notation, from now on $\mathbb{E}[\cdot]$ denotes the averages taken with respect to the underlying distribution from which $\Delta\boldsymbol{p}$ and $\boldsymbol{y}$ are sampled. In addition, we introduce an empirical measure $\langle\cdot\rangle$ to denote the averages with respect to the empirical sample under investigation. In particular, simple empirical estimators for the average price variation and the mean imbalance write

$$\hat{\boldsymbol{p}}_0 = \langle \Delta\boldsymbol{p} \rangle = \frac{1}{T}\sum_{t=1}^{T} \Delta\boldsymbol{p}^{(t)}, \tag{5.2}$$

$$\hat{\boldsymbol{y}}_0 = \langle \boldsymbol{y} \rangle = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{y}^{(t)}\,. \tag{5.3}$$

Henceforth, we consider that price changes and order flows are shifted by their empirical mean $\hat{\boldsymbol{p}}_0$ and $\hat{\boldsymbol{y}}_0$, and therefore we set $\hat{\boldsymbol{p}}_0 = \hat{\boldsymbol{y}}_0 = 0$. The corresponding covariances are then defined accordingly as

$$\hat{\boldsymbol{\Sigma}} = \langle \Delta\boldsymbol{p}\Delta\boldsymbol{p}^\top \rangle\,, \tag{5.4}$$

$$\hat{\boldsymbol{\Omega}}^{\mathrm{d}} = \langle \boldsymbol{y}\boldsymbol{y}^\top \rangle\,, \tag{5.5}$$

$$\hat{\boldsymbol{R}}^{\mathrm{d}} = \langle \Delta\boldsymbol{p}\boldsymbol{y}^\top \rangle\,. \tag{5.6}$$

As $T \to \infty$, the empirical averages converge to the actual means and covariances. Our goal is to show different recipes for the calibration of the model (5.1), all relying on the estimators above. We shall compare such estimators with the *Kyle estimator* $\hat{\boldsymbol{\Lambda}}_{\mathrm{Kyle}}$, which, as we shall see, displays several interesting properties. In order to evaluate the quality of different calibrations, we consider a quadratic loss χ^2 defined below.

Definition 5.1 (loss). *Given an empirical measure $\langle\cdot\rangle$, we define the* loss χ^2 *as the function*

$$\chi^2 = \frac{1}{2}\left\langle (\hat{\Delta}\boldsymbol{p} - \Delta\boldsymbol{p})^\top \boldsymbol{M} (\hat{\Delta}\boldsymbol{p} - \Delta\boldsymbol{p}) \right\rangle\,, \tag{5.7}$$

where $\hat{\Delta}\boldsymbol{p} = \hat{\boldsymbol{\Lambda}}\boldsymbol{y}$ is a linear predictor of the efficient price variation and $\boldsymbol{M}$ is a SPD matrix.

This definition of the loss implicitly implies that the calibration of the model is very similar to the task of the MM in the Kyle model: whereas in the latter the job of the MM is to forecast what the *fundamental* price variation is on the basis of the order flow imbalance, in this setting one is required to forecast the *efficient* (observed) price on the basis of the imbalance. Even though the two problems are different, they involve extremely similar equations, a fact that will allow us to leverage the results of the previous sections, valid in principle for a Kyle MM, in the calibration setting. Hence, under this definition of loss one is only required to find proxies for the price variations $\Delta\boldsymbol{p}$ and the dressed order imbalance $\boldsymbol{y}$, disregarding in principle the realizations of the underlying fundamental price $\boldsymbol{v}$.

5.2. Cross-impact estimators. Here we present three different cross-impact estimators and discuss their properties. The proofs of the propositions in this section are provided in Appendix B.

5.2.1. Maximum likelihood estimation. The simplest estimation recipe that one can construct in order to estimate the cross-impact matrix $\mathbf{\Lambda}$ is probably the MLE, $\hat{\mathbf{\Lambda}}_{\mathrm{MLE}}$, which is obtained by minimizing a quadratic loss.

Proposition 5.2. *The minimization of the loss χ^2 with respect to $\hat{\mathbf{\Lambda}}_{\mathrm{MLE}}$ yields*

$$\hat{\mathbf{\Lambda}}_{\mathrm{MLE}} = \hat{\boldsymbol{R}}^{\mathrm{d}}(\hat{\mathbf{\Omega}}^{\mathrm{d}})^{-1} , \tag{5.8}$$

independent of $\boldsymbol{M}$. The loss at the minimum is given by

$$\chi^2_{\mathrm{MLE}} = \frac{1}{2}\mathrm{tr}\left(\boldsymbol{M}(\hat{\mathbf{\Sigma}} - \hat{\mathbf{\Sigma}}_{\mathrm{MLE}})\right) , \tag{5.9}$$

where $\hat{\mathbf{\Sigma}}_{\mathrm{MLE}} = \hat{\boldsymbol{R}}^{\mathrm{d}}(\hat{\mathbf{\Omega}}^{\mathrm{d}})^{-1}(\hat{\boldsymbol{R}}^{\mathrm{d}})^{\top}$ is the portion of the covariance of fundamental price variations explained by the model (5.8).

Note that (5.8) is almost identical to (3.4), which has been obtained by enforcing the efficient price condition. This is no coincidence, as in a linear equilibrium setting with normally distributed variables the traded price is exactly a linear function of the order flow imbalance $\boldsymbol{y}$. There are though two important differences distinguishing the linear equilibrium of the Kyle model and the minimization of χ^2_{MLE}:

- An MM in the Kyle setting is actually performing a linear regression of the *fundamental* price, whereas here we are interested in regressing the *efficient* price. This justifies why (3.4) uses the *fundamental* price response $\boldsymbol{R}^d_v$, as opposed to (5.8), which uses the efficient price $\boldsymbol{R}^d$.
- In the Kyle setup case, the MM is aware that a part of the order imbalance comes from the IT who owns information about the fundamental price and is optimizing his strategy assuming that the MM uses an MLE. This leads to a quadratic equation for $\mathbf{\Lambda}$ (see (5.16) below) even in the linear equilibrium setting, whereas (5.8) above is a linear relation for $\hat{\mathbf{\Lambda}}_{\mathrm{MLE}}$.

Remark 2. One could have defined the multivariate Kyle model by replacing the efficient-price condition (see (2.2)) with a condition on the minimization of χ^2 without affecting the linear equilibrium. Moreover, this approach would do a better job at generalizing the model to more exotic settings. In fact, (A.5) provides a linear predictor only under the assumption of normality for prices and volumes: for other distributions, the relation linking $\boldsymbol{y}$ and $\boldsymbol{v}$ is in general nonlinear. Having an MM that minimizes a loss χ^2 is a reasonable assumption if one thinks that an MM without enough knowledge of the underlying distributions or computational power to build an efficient price should rely on linear regressions in order to estimate the latter.

The expression for $\hat{\mathbf{\Lambda}}_{\mathrm{MLE}}$ only contains the dressed empirical estimators $\hat{\boldsymbol{R}}^{\mathrm{d}}$ and $\hat{\mathbf{\Omega}}^{\mathrm{d}}$, allowing one to estimate the impact matrix from real data. In addition, such an estimator has the benefit of being very simple to implement, as it only requires the solution of a linear equation for $\hat{\mathbf{\Lambda}}_{\mathrm{MLE}}$. Unfortunately, this estimator lacks several properties that turn out to be very useful in cases of practical interest:
Symmetry: The matrix $\hat{\mathbf{\Lambda}}_{\mathrm{MLE}}$ is symmetric if and only if $\hat{\boldsymbol{R}}^{\mathrm{d}}$ and $\hat{\mathbf{\Omega}}^{\mathrm{d}}$ commute.
Positive definiteness: The matrix $\hat{\mathbf{\Lambda}}_{\mathrm{MLE}}$ can have negative eigenvalues (see [2]).

Consistency of correlations: In general, $[\hat{\mathbf{\Sigma}}, \hat{\mathbf{\Sigma}}_{\mathrm{MLE}}] \neq 0$. Hence, the price variations inferred by using the order flow imbalance do not share the eigenvectors of the real price variations, unless the response $\hat{\boldsymbol{R}}^{\mathrm{d}}(\hat{\mathbf{\Omega}}^{\mathrm{d}})^{-1}\hat{\boldsymbol{R}}^{\mathrm{d}}$ commutes with $\hat{\mathbf{\Sigma}}$.

The first two properties are extremely important in the calibration of cross-impact models, as shown in [1, 14]: to perform pricing within a cross-impact setup, the matrix $\mathbf{\Lambda}$ should be SPD in order to ensure the absence of price manipulation. This is not the case for an MLE, which is thus not suitable for practical purposes.

5.2.2. EigenLiquidity model. In order to cure the lack of symmetry of the MLE, the idea proposed in [12] is to construct an estimator of cross-impact $\hat{\mathbf{\Lambda}}_{\mathrm{ELM}}$ that is symmetric by construction, by enforcing the relation

$$[\hat{\mathbf{\Lambda}}_{\mathrm{ELM}}, \hat{\mathbf{\Sigma}}] = 0\,, \tag{5.10}$$

which means that the impact eigendirections coincide with the eigenportfolios (or principal components of the asset space). This prevents price manipulations that would be induced by an asymmetric of $\mathbf{\Lambda}$. Its calibration is explained in the next proposition.

Proposition 5.3. *Consider a pricing rule $\Delta\hat{\boldsymbol{p}} = \hat{\mathbf{\Lambda}}_{\mathrm{ELM}}\boldsymbol{y}$, in which we impose the commutation relation (5.10). Then, the MLE obtained under such a constraint takes the form*

$$\hat{\mathbf{\Lambda}}_{\mathrm{ELM}} = \sum_{a=1}^{n} \hat{\boldsymbol{s}}_a g_a \hat{\boldsymbol{s}}_a^{\top}\,, \tag{5.11}$$

where $\{\hat{\boldsymbol{s}}_a\}_{a=1}^n$ are the empirical eigenvectors of $\hat{\mathbf{\Sigma}}$, and where

$$g_a = \frac{\hat{\boldsymbol{s}}_a^{\top}\hat{\boldsymbol{R}}^{\mathrm{d}}\hat{\boldsymbol{s}}_a}{\hat{\boldsymbol{s}}_a^{\top}\hat{\mathbf{\Omega}}^{\mathrm{d}}\hat{\boldsymbol{s}}_a}\,. \tag{5.12}$$

Furthermore, the loss at the minimum is given by

$$\chi^2_{\mathrm{ELM}} = \frac{1}{2}\mathrm{tr}\left[M\left(\hat{\mathbf{\Sigma}} - \hat{\boldsymbol{R}}^{\mathrm{d}}\sum_{a=1}^{n}\hat{\boldsymbol{s}}_a\frac{\hat{\boldsymbol{s}}_a^{\top}\hat{\boldsymbol{R}}^{\mathrm{d}}\hat{\boldsymbol{s}}_a}{\hat{\boldsymbol{s}}_a^{\top}\hat{\mathbf{\Omega}}^{\mathrm{d}}\hat{\boldsymbol{s}}_a}\hat{\boldsymbol{s}}_a^{\top}\right)\right]\,. \tag{5.13}$$

In this case, the properties discussed above become the following:

Symmetry: The matrix $\hat{\mathbf{\Lambda}}_{\mathrm{ELM}}$ is symmetric by construction.

Positive definitenes: The matrix $\hat{\mathbf{\Lambda}}_{\mathrm{ELM}}$ can still have negative eigenvalues, although empirically the matrix $\hat{\mathbf{\Lambda}}_{\mathrm{ELM}}$ has been reported not to display eigenvalues significantly smaller than zero (see [12]).

Consistency of correlations: Similarly to the MLE case, the price variation covariances $\hat{\mathbf{\Sigma}}_{\mathrm{ELM}} = \hat{\mathbf{\Lambda}}_{\mathrm{ELM}}\hat{\mathbf{\Omega}}^{\mathrm{d}}\hat{\mathbf{\Lambda}}_{\mathrm{ELM}}$ do not generally commute with $\hat{\mathbf{\Sigma}}$. It is the case if and only if $[\hat{\mathbf{\Omega}}^{\mathrm{d}}, \hat{\mathbf{\Sigma}}] = 0$.

Summarizing, the price to pay in order to have a symmetric estimator is a larger loss function. Note that if $\hat{\boldsymbol{R}}^{\mathrm{d}}$ and $\hat{\mathbf{\Omega}}^{\mathrm{d}}$ commute with $\hat{\mathbf{\Sigma}}$, then $\hat{\mathbf{\Lambda}}_{\mathrm{ELM}} = (\hat{\boldsymbol{R}}^{\mathrm{d}})^2(\hat{\mathbf{\Omega}}^{\mathrm{d}})^{-1} = \hat{\mathbf{\Lambda}}_{\mathrm{MLE}}$ and therefore the loss is equal to the best possible $\chi^2_{\mathrm{ELM}} = \chi^2_{\mathrm{MLE}}$.

Remark 3 (estimators in one dimension). Since commutation and symmetry are granted in one dimension, imposing the condition (5.10) does not add any constraints to the minimization of the loss function and therefore $\hat{\lambda}_{\mathrm{ELM}} = \hat{\lambda}_{\mathrm{MLE}} = \frac{r^d}{\omega^d}$.

5.2.3. Kyle estimator. Now, let us provide some intuition on how to construct an impact estimator inspired by Kyle's model, $\hat{\mathbf{\Lambda}}_{\mathrm{Kyle}}$, by characterizing the $\mathbf{\Lambda}$ in the Kyle model from a different perspective. First, take the MM solution equation (3.4) (matching the MLE, equation (5.8)) and exchange the dressed order imbalancees and responses with those predicted in the linear equilibrium through equations (3.17) and (3.12):

$$\boldsymbol{R}_v^{\mathrm{d}} \to \frac{1}{2}\mathbf{\Sigma}_0\mathbf{\Lambda}^{-1}\,, \tag{5.14}$$

$$\mathbf{\Omega}^{\mathrm{d}} \to 2\mathbf{\Omega}\,. \tag{5.15}$$

Then, the linear regression of (3.4) becomes a quadratic matrix equation:

$$\mathbf{\Lambda} = \boldsymbol{R}_v^{\mathrm{d}}(\mathbf{\Omega}^{\mathrm{d}})^{-1} \to \mathbf{\Lambda} = \frac{1}{4}\mathbf{\Sigma}_0\mathbf{\Lambda}^{-1}\mathbf{\Omega}^{-1}\,. \tag{5.16}$$

After imposing symmetry and positive definiteness of $\mathbf{\Lambda}$, this is exactly the equation that leads to the expression of $\mathbf{\Lambda}$ obtained in the Kyle linear equilibrium (see Appendix A.4). By doing so, one ensures an efficient price variation process whose associated covariance is half that of the "true" price variations. As alluded to above, in the Kyle setup the MM is implicitly performing a linear regression and *knows that the IT is aware of his/her algorithm*, finally leading to (5.16). We now propose an estimator $\hat{\mathbf{\Lambda}}_{\mathrm{Kyle}}$ that exploits these ideas.

Proposition 5.4. *Let us consider an estimator $\hat{\Delta}\boldsymbol{p} := \hat{\mathbf{\Lambda}}_{\mathrm{Kyle}}\boldsymbol{y}$ such that*

- $\hat{\mathbf{\Lambda}}_{\mathrm{Kyle}}$ *is SPD,*
- *the empirical covariance of the associated efficient price variations $\hat{\mathbf{\Sigma}}_{\mathrm{Kyle}} = \langle\hat{\Delta}\boldsymbol{p}\hat{\Delta}\boldsymbol{p}^\top\rangle$ satisfies*

$$\hat{\mathbf{\Sigma}}_{\mathrm{Kyle}} = k^2\hat{\mathbf{\Sigma}}$$

with $k \in \mathbb{R}$. Then the unique $\hat{\mathbf{\Lambda}}_{\mathrm{Kyle}}$ satisfying these constraints is given by

$$\hat{\mathbf{\Lambda}}_{\mathrm{Kyle}} = k(\hat{\mathcal{R}}^{\mathrm{d}})^{-1}\sqrt{\hat{\mathcal{R}}^{\mathrm{d}}\hat{\mathbf{\Sigma}}\hat{\mathcal{L}}^{\mathrm{d}}}(\hat{\mathcal{L}}^{\mathrm{d}})^{-1}, \tag{5.17}$$

where $\hat{\mathbf{\Omega}}^{\mathrm{d}} = \hat{\mathcal{L}}^{\mathrm{d}}\hat{\mathcal{R}}^{\mathrm{d}}$ is a decomposition of $\hat{\mathbf{\Omega}}^{\mathrm{d}}$ such that $\hat{\mathcal{L}}^{\mathrm{d}} = (\hat{\mathcal{R}}^{\mathrm{d}})^\top$.

Moreover, the loss at the minimum is

$$\chi^2_{\mathrm{Kyle}} = \frac{1}{2}\mathrm{tr}\left[\boldsymbol{M}\left((1+k^2)\hat{\mathbf{\Sigma}} - 2k\hat{\boldsymbol{R}}^{\mathrm{d}}(\hat{\mathcal{R}}^{\mathrm{d}})^{-1}\sqrt{\hat{\mathcal{R}}^{\mathrm{d}}\hat{\mathbf{\Sigma}}\hat{\mathcal{L}}^{\mathrm{d}}}(\hat{\mathcal{L}}^{\mathrm{d}})^{-1}\right)\right] \tag{5.18}$$

and the value of k that minimizes the loss is given by

$$k^\star = \frac{\mathrm{tr}(\boldsymbol{M}\hat{\boldsymbol{R}}^{\mathrm{d}}(\hat{\mathcal{R}}^{\mathrm{d}})^{-1}\sqrt{\hat{\mathcal{R}}^{\mathrm{d}}\hat{\mathbf{\Sigma}}\hat{\mathcal{L}}^{\mathrm{d}}}(\hat{\mathcal{L}}^{\mathrm{d}})^{-1})}{\mathrm{tr}(\boldsymbol{M}\hat{\mathbf{\Sigma}})}\,.$$

Note that the complete analogy with the Kyle model is recovered for $k = 1$. In fact, if the sample dataset is drawn from an actual multivariate Kyle model with $\mathbf{\Sigma}_0/2 = \mathbf{\Sigma} = \hat{\mathbf{\Sigma}}$ and $2\mathbf{\Omega} = \mathbf{\Omega}^{\mathrm{d}} = \hat{\mathbf{\Omega}}^{\mathrm{d}}$ as described in the previous sections, one obtains

$$\mathbb{E}[\hat{\mathbf{\Lambda}}_{\mathrm{Kyle}}] \xrightarrow{T\to\infty} k\mathbf{\Lambda}\,.$$

The advantage of leaving k as a free parameter is that it allows one to increase or decrease the loss χ^2 of the Kyle estimator without affecting the eigenvectors of the predicted covariance. In fact, the relationship $2\mathbf{\Sigma} = \mathbf{\Sigma}_0$ arising in the Kyle model is thought not to be universal, so that it is more natural to think of it as a phenomenological parameter expressing the informational content of trades.

The adoption of $\hat{\mathbf{\Lambda}}_{\text{Kyle}}$ as an estimator of cross-impact has several interesting advantages with respect to more customary estimators such as the MLE $\hat{\mathbf{\Lambda}}_{\text{MLE}}$ or an impact estimator based on the EigenLiquidity Model $\hat{\mathbf{\Lambda}}_{\text{ELM}}$.

Symmetry and positive definiteness. The Kyle construction imposes a priori symmetry and PDness due to the request of optimality of the IT. In fact, the optimality condition (3.3) and the stability one (PDness of $\mathbf{\Lambda}$) are exactly the ones that a risk-neutral rational agent, as considered in [1, 14], would face when trying to optimize his profits.

Consistency of correlations. This construction allows one to recover the empirical correlations $\hat{\mathbf{\Sigma}}$ for $k = 1$. For *any value* of k one has $[\hat{\mathbf{\Sigma}}, \hat{\mathbf{\Sigma}}_{\text{Kyle}}] = 0$.

Loss. An important point concerns the loss function obtained by using $\hat{\mathbf{\Lambda}}_{\text{Kyle}}$. In general χ^2_{Kyle} is *larger* than that obtained by doing MLE, which minimizes χ^2 by design. Hence, the price to pay in order to have symmetry, positive semidefiniteness, and consistency of correlations is in general a worse fit of empirical data with respect to $\hat{\mathbf{\Lambda}}_{\text{MLE}}$ (as measured by the loss χ^2). However, if our data behaves as the Kyle model, i.e., if $\hat{\boldsymbol{R}}^{\text{d}} = \hat{\mathbf{\Sigma}}\hat{\mathbf{\Lambda}}^{-1}_{\text{Kyle}}$, then $k^\star = 1$ and

$$\chi^2_{\text{Kyle}} = 0 = \chi^2_{\text{MLE}} ,$$

implying that $\mathbf{\Lambda}$ is (*obviously!*) the best estimator for a market that reproduces the statistics of the multivariate Kyle model. Moreover, the loss at the minimum is zero because the *efficient* price variation $\Delta\boldsymbol{p}$ is completely determined by the imbalance. Indeed, the problem of regressing the *fundamental* price variation $\Delta\boldsymbol{v}$ would have yielded a nonzero loss at the minimum even for data sampled from an actual Kyle model due to the relation $\mathbf{\Sigma} < \mathbf{\Sigma}_0$.

Remark 4 (ELM and Kyle estimators for rank one $\hat{\mathbf{\Sigma}}$). As we showed in Example 8, when the price variation covariance is rank one $\hat{\mathbf{\Lambda}}_{\text{Kyle}}$ is proportional to $\hat{\mathbf{\Sigma}}$. On the other hand, $\hat{\mathbf{\Lambda}}_{\text{ELM}}$ has, by definition, the same eigenbasis as the price variations covariance. Therefore, if $\hat{\mathbf{\Sigma}}$ is rank one, then $\hat{\mathbf{\Lambda}}_{\text{ELM}}$ is proportional to $\hat{\mathbf{\Lambda}}_{\text{Kyle}}$ and $\hat{\mathbf{\Sigma}}$.

5.3. A comparison of recipes. In order to compare the performance of the estimators we compute observables that are directly related to the properties discused above:

- **Loss:** We compute the loss χ^2 as given in (5.7).
- **Asymmetry:** To quantify the degree of symmetry of the estimator we compute the norm of its antisymmetric part divided by the total norm:

$$\alpha = \frac{|\hat{\mathbf{\Lambda}} - \hat{\mathbf{\Lambda}}^\top|}{2|\hat{\mathbf{\Lambda}}|} .$$

  The value $\alpha = 0$ ($\alpha = 1$) means that the estimator is symmetric (antisymmetric).

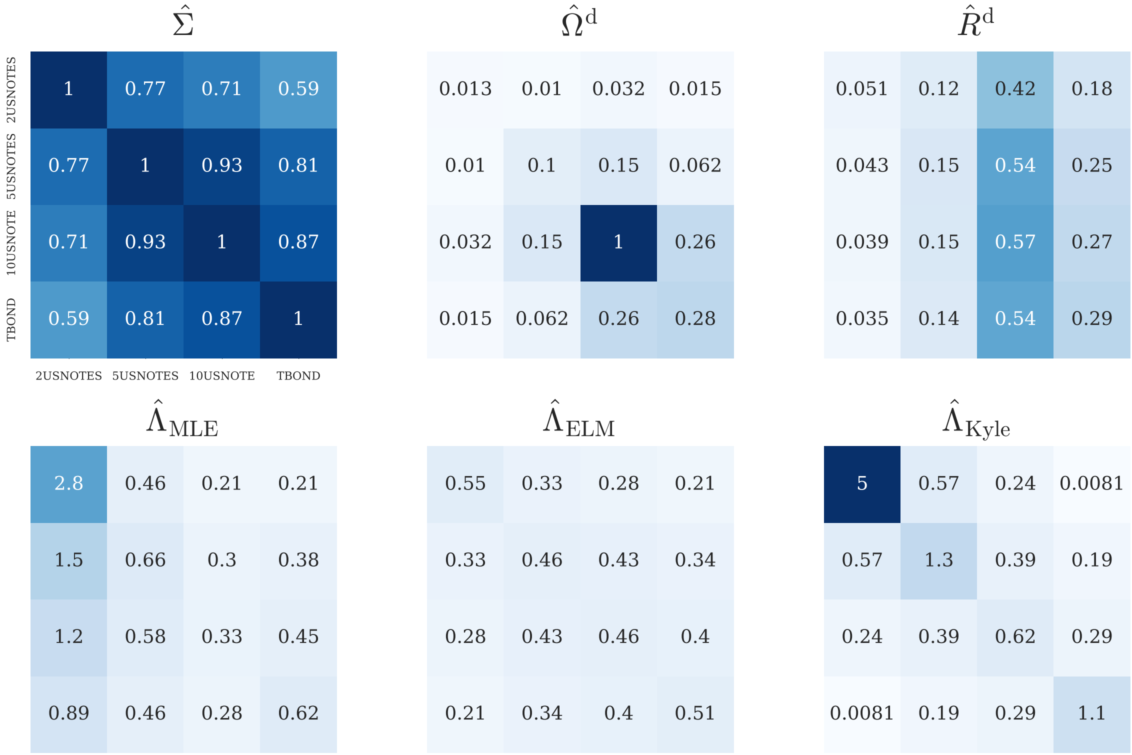

Figure 1. ***Averaged US Treasury Futures covariances and impact estimators for* 2016.** *Top from left to right: Daily averaged price variation covariances in untis of risk squared. Unitless daily averaged volume covariances rescaled with respect to the maximum. To obtain the original volumes* $\hat{\boldsymbol{\Omega}}^{\mathrm{d}}$ *has to be multiplied by* \$2242^2. *Market response rescaled with respect to the maximal volume in units of risk. To obtain the original volumes* $\hat{\boldsymbol{R}}^{\mathrm{d}}$ *has to be multiplied by* \$$\sqrt{2242}$. *Bottom from left to right: maximum likelihood impact estimator, ELM based impact estimator and Kyle model based impact estimator in units of risk. To obtain the estimators in the right units they have to be divided by* \$$\sqrt{2242}$.

- **Positive definiteness:** We compute the lowest real part of the spectrum of the estimator:

$$\lambda^\star = \min_i \Re(\hat{\lambda}_i) \ ,$$

where $\{\hat{\lambda}_i\}_{1\geq i\geq n}$ are the eigenvalues of $\hat{\boldsymbol{\Lambda}}$. Positive definiteness is equivalent to $\lambda^\star > 0$.
- **Consistency of correlations:** We compute the norm of the commutator of $\hat{\boldsymbol{\Sigma}}$ and the covariances of the estimated price variations $\hat{\boldsymbol{\Sigma}}_{\mathrm{est}}$:

$$\kappa = \frac{|\hat{\boldsymbol{\Sigma}}_{\mathrm{est}}\hat{\boldsymbol{\Sigma}} - \hat{\boldsymbol{\Sigma}}\hat{\boldsymbol{\Sigma}}_{\mathrm{est}}|}{|\hat{\boldsymbol{\Sigma}}||\hat{\boldsymbol{\Sigma}}_{\mathrm{est}}|} \ .$$

Correlations are consistent if $\kappa = 0$.

5.3.1. Application to real data. In Figure 1 we show the covariances and estimators for the 2-year, 5-year, 10-year, and 30-year tenors of the U.S. Treasury Futures traded in the Chicago Board of Trade (see B.5 for details). In Table 1 we show the values of the observables

Table 1

Values of the observables for the three estimators for U.S. Treasury Futures and Bonds.

Estimator	Loss (χ^2)	Commutator (κ)	PDness ($\lambda^\star$)	Asymmetry (α)
$\hat{\mathbf{\Lambda}}_{\text{MLE}}$	1.20	0.038	0.031	0.31
$\hat{\mathbf{\Lambda}}_{\text{ELM}}$	1.26	0.129	0.017	0
$\hat{\mathbf{\Lambda}}_{\text{Kyle}}$	1.32	0	0.386	0

Table 2

Average values of the observables for the three estimators for the six pairs of U.S. Treasury Futures.

Estimator	Loss (χ^2)	Commutator (κ)	PDness ($\lambda^\star$)	Asymmetry (α)
$\hat{\mathbf{\Lambda}}_{\text{MLE}}$	0.668	0.033	0.318	0.186
$\hat{\mathbf{\Lambda}}_{\text{ELM}}$	0.691	0.108	0.231	0
$\hat{\mathbf{\Lambda}}_{\text{Kyle}}$	0.718	0	0.952	0

corresponding to the 4×4 system. In Table 2 we show the values of the same observables but averaged over the six possible combinations of 2×2 systems. As expected, the Kyle procedure fares slightly worse for the loss function but is much better than other procedures on all other counts.

5.3.2. Synthetic data. To further explore the effects of price variation and liquidity correlations, in the next two sections we compute the values of the observables described above on the six sets of synthetic 2×2 covariance matrices with specific price variation covariances $(\hat{\mathbf{\Sigma}}_\rho, \hat{\mathbf{\Omega}}^{\text{d}}_\rho, \hat{\boldsymbol{R}}^{\text{d}}_\rho)$ and specific volume covariances $(\hat{\mathbf{\Sigma}}_\epsilon, \hat{\mathbf{\Omega}}^{\text{d}}_\epsilon, \hat{\boldsymbol{R}}^{\text{d}}_\epsilon)$, respectively, fabricated by modifying the data for U.S. Treasury Futures as explained in B.5 (see Table 2). The idea behind this construction is to provide a synthetic but realistic set of matrices that are parametrized by either a liquidity parameter ϵ or a correlation parameter ρ. By varying those parameters, one can extrapolate from the reality (recovered for specific values of ϵ and/or ρ) and a fictitious world in which assets can be made more or less liquid with respect to reality by varying the knob ϵ, and more or less correlated by varying the parameter ρ.

Extreme illiquidity. In order to explore the effect of extreme heterogeneous liquidities we rescale the 2×2 covariances as

$$\hat{\mathbf{\Omega}}^{\text{d}} \mapsto \hat{\mathbf{\Omega}}^{\text{d}}_\epsilon = \begin{pmatrix} 1 & \sqrt{\epsilon}\,\hat{\mathbf{\Omega}}^{\text{d}}_{12} \\ \sqrt{\epsilon}\,\hat{\mathbf{\Omega}}^{\text{d}}_{12} & \epsilon \end{pmatrix} .$$

In Figure 2 we show the value of each of the observables averaged over the six sets of covariances for ϵ ranging from 0 to 1. Regarding the loss, the estimator that better performs is $\hat{\mathbf{\Lambda}}_{\text{MLE}}$, regardless of ϵ as expected. It is interesting to observe that χ^2_{MLE} does not depend on ϵ. The reason is that, by construction, the estimated price variation covariances are invariant under changes in the volume variances:[3]

$$\hat{\mathbf{\Sigma}}_{\text{MLE}} = \hat{\boldsymbol{R}}^{\text{d}}_\epsilon (\hat{\mathbf{\Omega}}^{\text{d}}_\epsilon)^{-1} (\hat{\boldsymbol{R}}^{\text{d}}_\epsilon)^\top = \hat{\boldsymbol{R}}^{\text{d}} (\hat{\mathbf{\Omega}}^{\text{d}})^{-1} (\hat{\boldsymbol{R}}^{\text{d}})^\top .$$

The loss of the other two estimators depends on ϵ. For small values of ϵ, χ^2_{Kyle} is lower than χ^2_{ELM}, suggesting that $\hat{\mathbf{\Lambda}}_{\text{Kyle}}$ might be better in the case of markets with heterogeneous

[3]This is not a particular property of systems in two dimensions; it holds for all dimensions.

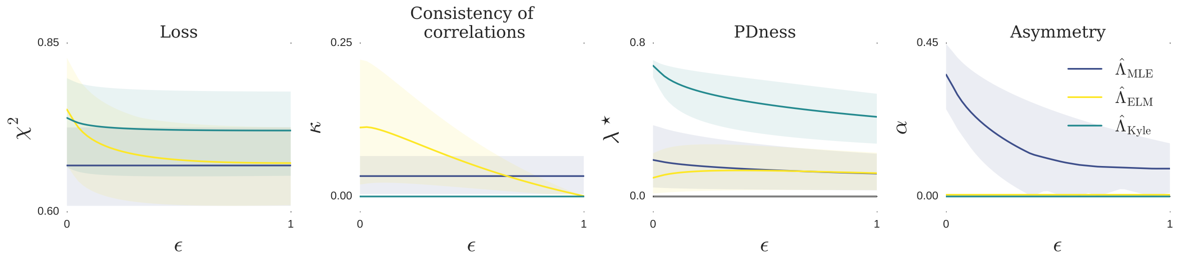

Figure 2. ***Effect of liquidity.*** *Averaged values (solid lines) and intervals between the maximum and the minimum (shaded area) of the observables for* ϵ *ranging from* 0 *to* 1.

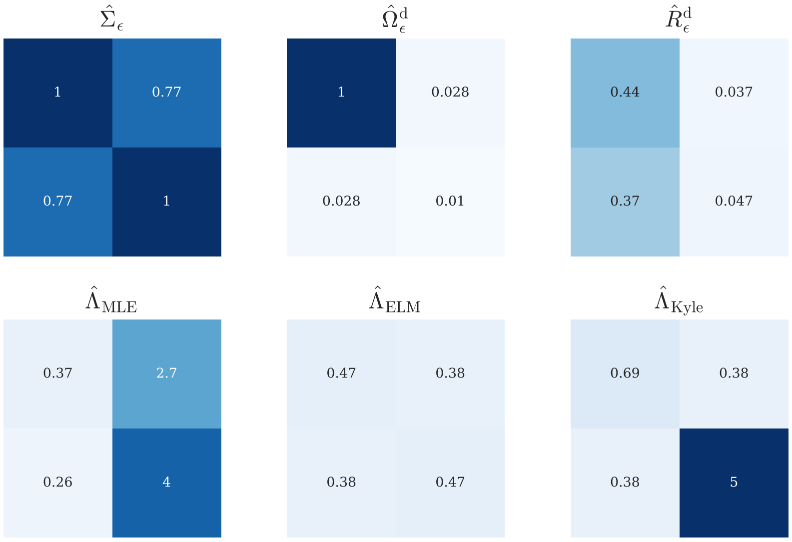

Figure 3. ***Extreme illiquidity.*** *Covariance matrices and estimators with an illiquid asset. The value of* $\rho = 0.77$ *is the empirical correlation of the representative pair* 2*USNOTES*-5*USNOTES, whereas the value of* $\epsilon = 0.01$ *is obtained by rescaling the empirical responses and volume covariances of the same pair.*

volatilities. For $\epsilon \gtrsim 0.2$ the loss distributions for the three estimators have a lot of overlap and no one is significantly better.

As explained above, κ_{MLE} does not depend on ϵ and $\kappa_{\mathrm{Kyle}} = 0$, by construction. On the other hand, the limit κ_{ELM} decreases to 0 as $\epsilon \to 1$ because in the limit $\epsilon = 1$, $\hat{\boldsymbol{\Omega}}^{\mathrm{d}}_\epsilon$ and $\hat{\boldsymbol{\Sigma}}_\epsilon$ commute. All estimators are PD, and the only nonsymmetric estimator is $\hat{\boldsymbol{\Lambda}}_{\mathrm{MLE}}$, whose antisymmetric part decreases as ϵ increases (even though it never becomes fully symmetric). The other two are symmetric by construction.

Reminiscent to what we showed in Example 6, in Figure 3 we display the covariance matrices and the three corresponding impact estimators in a case where one of the assets' liquidity is much smaller than the other one. Note that due to the lack of liquidity of the second asset the second column of $\hat{\boldsymbol{R}}^{\mathrm{d}}_\epsilon$ is considerably smaller than the first one while price

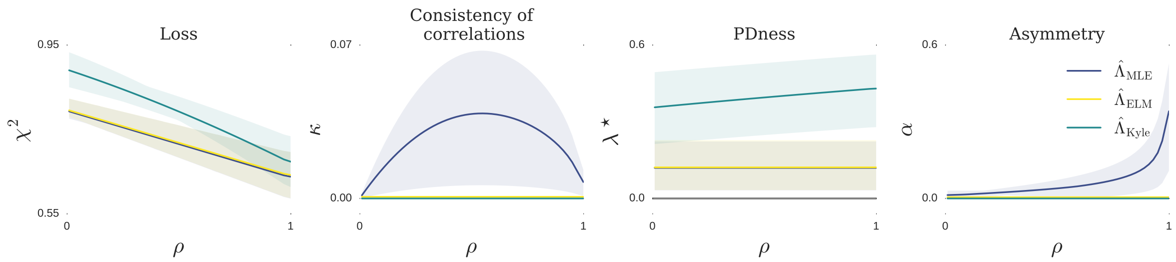

Figure 4. ***Effect of price variation correlations*** *Averaged values (solid lines) and intervals between the maximum and the minimum (shaded area) of the observables for ρ ranging from 0 to 1.*

covariances are not changed. Let us analyze heuristically the estimators obtained:

- $\hat{\mathbf{\Lambda}}_{\text{MLE}}$: Trading the second asset will have a strong impact on its price and a significant impact on the price of the first asset (due to the nonnegligible price correlations). Trading the first asset will have little impact on both prices.
- $\hat{\mathbf{\Lambda}}_{\text{ELM}}$: In this case the impact estimator predicts that trading either asset will have a similar impact.
- $\hat{\mathbf{\Lambda}}_{\text{Kyle}}$: The main difference between the predictions of the Kyle estimator and $\hat{\mathbf{\Lambda}}_{\text{MLE}}$ is that in this case trading the second asset will strongly modify its price but it will have very little impact on the first assets' price. As already emphasized, this is how efficient markets should work: liquid instruments should not be affected by the order flow on correlated illiquid instruments; otherwise price manipulation strategies would be possible.

Extreme correlations. We now look at a linear combinations of two of the assets with price variation covariances:

$$\hat{\mathbf{\Sigma}} \mapsto \hat{\mathbf{\Sigma}}_\rho = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} .$$

In Figure 4 we show the value of each of the observables averaged over the six sets of covariances for ϵ ranging from 0 to 1. The loss of all estimators decreases as $\rho \to 1$ and as in the previous case, the distributions have a strong overlap, suggesting again that no one estimator is significantly better at minimizing the loss.

By construction $\kappa_{\text{Kyle=0}}$, and in this case $\kappa_{\text{ELM}} = 0$ because $\hat{\mathbf{\Omega}}^{\text{d}}_\rho$ commutes with $\hat{\mathbf{\Sigma}}_\rho$ (it would not be the case if the volume variances were not the same). Also $\kappa_{\text{MLE}} = 0$ for $\rho = 0$ because in that case $\hat{\mathbf{\Sigma}}_\rho = \mathbb{I}$, but for $\rho > 0$, $\kappa_{\text{MLE}} > 0$. The distributions of $\lambda^\star_{\text{ELM}}$ and $\lambda^\star_{\text{MLE}}$ are essentially the same. $\lambda^\star_{\text{Kyle}}$ is significantly larger than the other two for all values of ρ. Asymmetries in $\hat{\mathbf{\Lambda}}_{\text{MLE}}$ are amplified as $\rho \to 1$ because in that limit $\hat{\mathbf{\Omega}}^{\text{d}}_\rho$ becomes singular.

Again, following Example 7 in Figure 5 we show the covariance matrices and the three corresponding impact estimators for a case where the price variations of the two assets are strongly correlated. Let us analyze the estimators obtained:

- $\hat{\mathbf{\Lambda}}_{\text{MLE}}$: Small arbitrary fluctuations in $\hat{\boldsymbol{R}}^{\text{d}}_\rho$ are amplified due to the strong volume correlations that result from the transformation. In the example shown in the figure the impact on the second asset is much larger than on the first one.

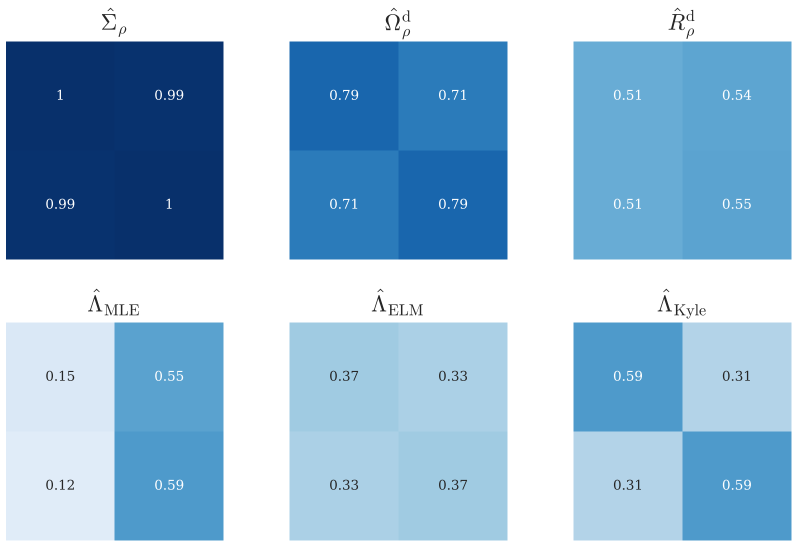

Figure 5. ***Strong price variation correlations.*** *Covariance matrices and estimators with strongly correlated price variations* ($\rho = 0.99$).

- $\hat{\mathbf{\Lambda}}_{\mathrm{ELM}}$: Self impact and cross-impact are comparable. Furthermore, self-impact is very similar for both assets.
- $\hat{\mathbf{\Lambda}}_{\mathrm{Kyle}}$: Self-impact is very similar for both assets and stronger than cross-impact.

6. Conclusion. In the present work we studied a multivariate Kyle model that proved to be a very interesting setting to understand the fundamental mechanisms underlying cross-impact. Perhaps more importantly, the multivariate Kyle model suggests a practical recipe to extract a consistent cross-impact matrix structure from empirical data—a point that does not seem to have been emphasized before but that becomes crucial when dealing with present day large dimensional data. We revisited a special case of the Caballé–Krishnan solution at equilibrium and proved the unicity of the symmetric solution. We provided an interpretation of the results with regard to the eigen-modes of returns and volumes covariances. We discussed the implications of the model for pricing, with a particular focus on the SPD property of the impact matrix. We presented the implications for cross-impact regression of empirical data and showed that cross-awareness (or in lesser terms "I know that you know") can be used as a powerful regularizer with a small loss in predictive power. We confronted our results with previous empirical cross-impact analyses and identified limiting regimes in which the results in [2, 12] are reproduced.

From a complementary point of view, our results can be seen as proxies for the behavior of an idealized market in which prices fully reflect the information encoded in the order flow: for a manipulation-free market, prices of liquid securities should be insensitive to trading of strongly illiquid instruments (see Example 6), whereas prices of strongly correlated instruments should be insensitive to how they are individually traded (Example 7). Measuring how much real markets violate these principles would be an interesting empirical application of our results, which could be used to assist regulators in order to assess the vulnerability of a market to correlated trading activity.

In spite of all the "good" properties of the Kyle cross-impact estimator, one should keep in mind that many important aspects are left out of the Kyle framework that may play a crucial role in practice. First, the empirical order flow is strongly autocorrelated in time, which must induce a nontrivial lag dependence of the impact function—as found in [2]. Extending the present theory to predict a lag-dependent impact matrix $\boldsymbol{\Lambda}(\tau)$ would be extremely useful. Second, impact is nonlinear in traded quantities: it is now well accepted that the impact of a metaorder has a square-root dependence on volume. How this single asset square-root law generalizes in a multiasset context is, as far as we are aware, a completely open problem. Finally, it would be interesting to extend to the multivariate case to some recent extensions of the Kyle model that account for the inventory risk of the MM [6]. We hope to visit some of these questions in future work.

Appendix A. Solution of the multivariate Kyle model. The proof of Theorem 3.5 is split into different results. First, the fact that the equilibrium is linear is a consequence of Propositions 3.2 and 3.3. Before finding the explicit form of the equilibrium we need to show that $\boldsymbol{\Lambda}$ is symmetric. This is a consequence of the second order condition on the minimization of the IT's strategy (Proposition 3.2) and is proven in Proposition 3.4. Finally, the explicit form and proof of uniqueness of $\boldsymbol{\Lambda}$ is given in Proposition A.2

A.1. Optimality of the informed trader.

Proof of Proposition 3.2. Here, we show the consequences of the first and second order conditions arising from the requirement that the IT is maximizing locally the average utility given $\boldsymbol{v}$, which can be written as

$$\begin{aligned}\mathbb{E}[\mathcal{U}_{IT}|\boldsymbol{v}] &= -\mathbb{E}[\boldsymbol{x}^\top(\boldsymbol{p}-\boldsymbol{v})|\boldsymbol{v}]\\ &= -\mathbb{E}[\boldsymbol{x}^\top(\boldsymbol{\mu}+\boldsymbol{\Lambda}\boldsymbol{y}-\boldsymbol{v})|\boldsymbol{v}]\\ &= -\boldsymbol{x}^\top\boldsymbol{\mu}-\boldsymbol{x}^\top\boldsymbol{\Lambda}\boldsymbol{x}+\boldsymbol{x}^\top\boldsymbol{v}.\end{aligned}$$

In order to find the optimal strategy, we maximize the utility with respect to $\boldsymbol{x}$. The first and seccond order derivatives are

$$\frac{\partial\mathbb{E}[\mathcal{U}_{IT}|\boldsymbol{v}]}{\partial\boldsymbol{x}} = -\boldsymbol{\mu}-2\boldsymbol{\Lambda}_{\mathrm{S}}\boldsymbol{x}+\boldsymbol{v}=0, \tag{A.1}$$

$$\frac{\partial^2\mathbb{E}[\mathcal{U}_{IT}|\boldsymbol{v}]}{\partial\boldsymbol{x}\partial\boldsymbol{x}^\top} = -2\boldsymbol{\Lambda}_{\mathrm{S}}, \tag{A.2}$$

where $\boldsymbol{X}_{\mathrm{S}}$ ($\boldsymbol{X}_{\mathrm{A}}$) denotes the symmetric (antisymmetric) part of $\boldsymbol{X}$.

Since $\mathcal{U}_{\mathrm{IT}}$ is quadratic on $\boldsymbol{x}$, the profit maximization condition implies that the second derivative (A.2) has to be strictly negative definite, which in turn implies that both $\boldsymbol{\Lambda}$ and $\boldsymbol{\Lambda}_{\mathrm{S}}$ have to be PD. Solving (A.1) for $\boldsymbol{x}$ leads to

$$\boldsymbol{x} = \boldsymbol{\alpha}+\boldsymbol{B}\boldsymbol{v}, \tag{A.3}$$

where

$$\begin{aligned} \boldsymbol{\alpha} &= -\boldsymbol{B}\boldsymbol{\mu}, \\ \boldsymbol{B} &= \frac{1}{2}\boldsymbol{\Lambda}_{\mathrm{S}}^{-1}. \end{aligned} \tag{A.4}$$

Since $\boldsymbol{\Lambda}_{\mathrm{S}}$ has to be PD, it is invertible and $\boldsymbol{B}$ and $\boldsymbol{\alpha}$ are well defined. ■

As we will show in Appendix A.3, the PDness of $\boldsymbol{\Lambda}$ together with the efficient-price condition (discussed in Appendix A.2) implies that $\boldsymbol{\Lambda}$ is symmetric.

A.2. Optimality of the market maker. In this appendix we will state the conditions stemming from the requirement that the MM fixes a pricing rule $\boldsymbol{p} = \boldsymbol{\Lambda}\boldsymbol{y} + \boldsymbol{\mu}$ such that $\boldsymbol{p} = \mathbb{E}[\boldsymbol{v}|\boldsymbol{y}]$ is an efficient price.

Proof of Proposition 3.3. Exploiting the Gaussian nature of $\boldsymbol{v}$ and $\boldsymbol{u}$ and their independence, the efficient price as given in (2.2) can be expanded as [11, p. 269]

$$\begin{aligned} \boldsymbol{p} = \mathbb{E}[\boldsymbol{v}|\boldsymbol{y}] &= \mathbb{E}[\boldsymbol{v}] + \mathbb{C}[\boldsymbol{v},\boldsymbol{y}]\mathbb{C}[\boldsymbol{y},\boldsymbol{y}]^{-1}(\boldsymbol{y} - \mathbb{E}[\boldsymbol{y}]) \\ &= \boldsymbol{p}_0 + \boldsymbol{R}_v^{\mathrm{d}}(\boldsymbol{\Omega}^{\mathrm{d}})^{-1}(\boldsymbol{y} - \boldsymbol{y}_0) \\ &= \boldsymbol{\mu} + \boldsymbol{\Lambda}\boldsymbol{y}\,, \end{aligned} \tag{A.5}$$

where

$$\boldsymbol{\mu} = \boldsymbol{p}_0 - \boldsymbol{\Lambda}\boldsymbol{y}_0\,, \tag{A.6}$$

$$\boldsymbol{\Lambda} = \boldsymbol{R}_v^{\mathrm{d}}(\boldsymbol{\Omega}^{\mathrm{d}})^{-1}. \tag{A.7}$$

■

This analysis provides an explicit form for the MM's strategy and it might look as if the MM is doing a linear regression. However, without knowing the strategy of the IT one cannot estimate $\boldsymbol{R}_v^{\mathrm{d}}$ nor $(\boldsymbol{\Omega}^{\mathrm{d}})^{-1}$, meaning that the MM has to include information about $\boldsymbol{x}$ in his regression, making it a quadratic problem rather than a linear one. This point is discussed in more detail in section 5.2.1.

A.3. Symmetry of the linear equilibrium. In order to find the linear equilibrium described in Theorem 3.5, we first have to prove the Proposition 3.4. To do so we will make use of a useful lemma [8, Theorem 7.2.6] that will be used in the following derivations.

Lemma A.1 (positive definite square-root). *Consider a matrix $\boldsymbol{Y}$ that is symmetric and positive semidefinite. Then, there exists a unique symmetric, positive semidefinite matrix $\boldsymbol{X} = \sqrt{\boldsymbol{Y}}$ such that $\boldsymbol{X}\boldsymbol{X} = \boldsymbol{Y}$.*

Proof of Proposition 3.4. Let us start by plugging $\mathbb{E}[\boldsymbol{y}] = \boldsymbol{y}_0 = \frac{1}{2}\boldsymbol{\Lambda}_{\mathrm{S}}^{-1}(\boldsymbol{p}_0 - \boldsymbol{\mu})$ (derived from (3.3)) into (3.4). After a bit of algebra we obtain

$$(\boldsymbol{p}_0 - \boldsymbol{\mu})\left(\mathbb{I} - \frac{1}{2}\boldsymbol{\Lambda}\boldsymbol{\Lambda}_{\mathrm{S}}^{-1}\right) = 0, \tag{A.8}$$

thus $\boldsymbol{\mu} = \boldsymbol{p}_0$ and $\boldsymbol{x} = \frac{1}{2}\boldsymbol{\Lambda}_{\mathrm{S}}^{-1}(\boldsymbol{v} - \boldsymbol{p}_0)$, which allows us to compute the following quantities:

$$\boldsymbol{y}_0 = 0, \tag{A.9}$$

$$\boldsymbol{\Omega}^{\mathrm{d}} = \frac{1}{4}\boldsymbol{\Lambda}_{\mathrm{S}}^{-1}\boldsymbol{\Sigma}_0\boldsymbol{\Lambda}_{\mathrm{S}}^{-1} + \boldsymbol{\Omega}, \tag{A.10}$$

$$\boldsymbol{R}_v^{\mathrm{d}} = \frac{1}{2}\boldsymbol{\Sigma}_0\boldsymbol{\Lambda}_{\mathrm{S}}^{-1}. \tag{A.11}$$

By plugging the results above into (3.4) one obtains a quadratic equation:

$$(\boldsymbol{\Lambda}\boldsymbol{\Lambda}_{\mathrm{S}}^{-1} - 2\mathbb{I})\frac{1}{4}\boldsymbol{\Sigma}_0 + \boldsymbol{\Lambda}\boldsymbol{\Omega}\boldsymbol{\Lambda}_{\mathrm{S}} = 0. \tag{A.12}$$

Expanding $\boldsymbol{\Lambda}$ as $(\boldsymbol{\Lambda}_{\mathrm{A}} + \boldsymbol{\Lambda}_{\mathrm{S}})$ and doing some algebra from (A.12) we obtain

$$\boldsymbol{\Lambda}_{\mathrm{A}} = (\mathbb{I} - \boldsymbol{\Gamma})(\mathbb{I} + \boldsymbol{\Gamma})^{-1}\boldsymbol{\Lambda}_{\mathrm{S}}, \tag{A.13}$$

where $\boldsymbol{\Gamma} = 4\boldsymbol{\Lambda}_{\mathrm{S}}\boldsymbol{\Omega}\boldsymbol{\Lambda}_{\mathrm{S}}\boldsymbol{\Sigma}_0^{-1}$. Now, since $\boldsymbol{\Lambda}_{\mathrm{A}} = \frac{1}{2}(\boldsymbol{\Lambda} - \boldsymbol{\Lambda}^{\top})$, we have $\boldsymbol{\Lambda}_{\mathrm{A}} = -\boldsymbol{\Lambda}_{\mathrm{A}}^{\top}$, and therefore

$$\begin{aligned}
&(\mathbb{I} - \boldsymbol{\Gamma})(\mathbb{I} + \boldsymbol{\Gamma})^{-1}\boldsymbol{\Lambda}_{\mathrm{S}} = -\boldsymbol{\Lambda}_{\mathrm{S}}(\mathbb{I} + \boldsymbol{\Gamma}^{\top})^{-1}(\mathbb{I} - \boldsymbol{\Gamma}^{\top}),\\
&(\mathbb{I} + \boldsymbol{\Gamma}^{\top})\boldsymbol{\Lambda}_{\mathrm{S}}^{-1}(\mathbb{I} - \boldsymbol{\Gamma}) = -(\mathbb{I} - \boldsymbol{\Gamma}^{\top})\boldsymbol{\Lambda}_{\mathrm{S}}^{-1}(\mathbb{I} + \boldsymbol{\Gamma}),\\
&\sqrt{\boldsymbol{\Lambda}_{\mathrm{S}}}\boldsymbol{\Gamma}^{\top}\boldsymbol{\Lambda}_{\mathrm{S}}^{-1}\boldsymbol{\Gamma}\sqrt{\boldsymbol{\Lambda}_{\mathrm{S}}} = \mathbb{I}.
\end{aligned}$$

Substituting back $\boldsymbol{\Gamma}$ we get

$$\boldsymbol{Y}^2 = \boldsymbol{Z}^2,$$

where $\boldsymbol{Y} = \sqrt{\boldsymbol{\Lambda}_{\mathrm{S}}}\boldsymbol{\Omega}\sqrt{\boldsymbol{\Lambda}_{\mathrm{S}}}$ and $\boldsymbol{Z} = \sqrt{\boldsymbol{\Lambda}_{\mathrm{S}}^{-1}\frac{1}{4}\boldsymbol{\Sigma}_0\sqrt{\boldsymbol{\Lambda}_{\mathrm{S}}^{-1}}}$. Since $\boldsymbol{\Lambda}_{\mathrm{S}}$ is PD, $\boldsymbol{Y}^2$ is also SPD and therefore $\boldsymbol{Y} = \boldsymbol{Z}$ is the unique PD square root of $\boldsymbol{Y}^2$, which leads to

$$\frac{1}{4}\boldsymbol{\Sigma}_0 = \boldsymbol{\Lambda}_{\mathrm{S}}\boldsymbol{\Omega}\boldsymbol{\Lambda}_{\mathrm{S}}, \tag{A.14}$$

and substituting these results into (A.13) reveals that $\boldsymbol{\Lambda}_{\mathrm{A}} = 0$. ■

With (A.14) we prove that the profit maximization condition on the IT's cost (i.e., $\boldsymbol{\Lambda}$ has to be PD) results in $\boldsymbol{\Lambda}_{\mathrm{A}} = 0$ and $\boldsymbol{\Lambda} = \boldsymbol{\Lambda}_{\mathrm{S}}$. If $\boldsymbol{\Lambda}_{\mathrm{S}}$ was not PD, $\boldsymbol{Y}$ and $\boldsymbol{Z}$ could have different roots and we would not be able to establish the relationship $\boldsymbol{Y} = \boldsymbol{Z}$. In that case we would find many solutions, most of them being saddle points for $\mathbb{E}[\mathcal{U}_{IT}]$.

A.4. Explicit form of the linear equilibrium. The last step needed in order to prove the existence and uniqueness of the linear equilibrium in Theorem 3.5 is to prove that (A.14) admits a unique symmetric solution, as made explicit below.

Proposition A.2. *The unique symmetric PD matrix $\boldsymbol{\Lambda}$ satisfying (A.14) is*

$$\boldsymbol{\Lambda} = \frac{1}{2}\mathcal{R}^{-1}\sqrt{\mathcal{R}\boldsymbol{\Sigma}_0\mathcal{L}}\mathcal{L}^{-1}.$$

Proof. *Explicit form of* $\mathbf{\Lambda}$*:* The solution to (A.14) has to be a $\mathbf{\Lambda}$ of the form

$$\mathbf{\Lambda} = \frac{1}{2}\mathcal{G}\boldsymbol{O}\mathcal{L}^{-1}\,, \tag{A.15}$$

where $\mathcal{G}$ is any factorization of $\mathbf{\Sigma}_0$ of the form $\mathcal{G}\mathcal{D}$ with $\mathcal{G}^\top = \mathcal{D}$. One is then left with the equation $\boldsymbol{O}\boldsymbol{O}^\top = \mathbb{I}$. Such a equation is solved by *any* matrix $\boldsymbol{O}$ belonging to the orthogonal group $O(n)$, which is specified by $n(n-1)/2$ parameters. It is only by self-consistently imposing symmetry of $\mathbf{\Lambda}$ that one finds the solution for $\boldsymbol{O}$,

$$\begin{aligned}\boldsymbol{O} &= \mathcal{G}^{-1}\mathcal{R}^{-1}\sqrt{\mathcal{R}\mathbf{\Sigma}_0\mathcal{L}}\\ &= (\mathcal{R}\mathcal{G})^{-1}\sqrt{(\mathcal{R}\mathcal{G})(\mathcal{R}\mathcal{G})^\top}.\end{aligned} \tag{A.16}$$

Introducing this value of $\boldsymbol{O}$ in (A.15) we obtain the desired result.

Uniqueness: The decomposition $\mathbf{\Omega} = \mathcal{L}\mathcal{R}$ is not unique. Indeed, multiplying $\mathcal{L}$ by an arbitrary rotation $\boldsymbol{O}_\Omega$ yields another possible decomposition. However, the final value of $\mathbf{\Lambda}$ does not depend on $\boldsymbol{O}_\Omega$. To prove this consider the value of $\mathbf{\Lambda}$ using the decomposition $\mathbf{\Omega} = \mathcal{L}\boldsymbol{O}_\Omega\boldsymbol{O}_\Omega^\top\mathcal{R}$:

$$\begin{aligned}\mathbf{\Lambda} &= \frac{1}{2}(\boldsymbol{O}_\Omega^\top\mathcal{R})^{-1}\sqrt{\boldsymbol{O}_\Omega^\top\mathcal{R}\mathbf{\Sigma}_0\mathcal{L}\boldsymbol{O}_\Omega}(\mathcal{L}\boldsymbol{O}_\Omega)^{-1}\\ &= \frac{1}{2}\mathcal{R}^{-1}\boldsymbol{O}_\Omega\boldsymbol{O}_\Omega^\top\sqrt{\mathcal{R}\mathbf{\Sigma}_0\mathcal{L}}\boldsymbol{O}_\Omega\boldsymbol{O}_\Omega^\top\mathcal{L}^{-1}\\ &= \frac{1}{2}\mathcal{R}^{-1}\sqrt{\mathcal{R}\mathbf{\Sigma}_0\mathcal{L}}\mathcal{L}^{-1},\end{aligned}$$

where we have used the fact that if $\sqrt{\boldsymbol{Y}} = \boldsymbol{X}$, then $\sqrt{\boldsymbol{O}^\top\boldsymbol{Y}\boldsymbol{O}} = \boldsymbol{O}^\top\boldsymbol{X}\boldsymbol{O}$. Since the argument of the square root is SPD by construction the unique value of $\mathbf{\Lambda}$ is obtained by choosing the unique SPD root. ∎

Appendix B. Estimation.

B.1. The loss χ^2. First, we want to relate the loss (5.7) with the empirical mean and covariances defined in section 5. Starting from the definition (5.7), one can write

$$\begin{aligned}\chi^2 &= \frac{1}{2}\langle(\hat{\mathbf{\Lambda}}\boldsymbol{y} - \Delta\boldsymbol{p})^\top\boldsymbol{M}(\hat{\mathbf{\Lambda}}\boldsymbol{y} - \Delta\boldsymbol{p})\rangle\,,\\ &= \frac{1}{2}\langle\boldsymbol{y}^\top\hat{\mathbf{\Lambda}}^\top\boldsymbol{M}\hat{\mathbf{\Lambda}}\boldsymbol{y}\rangle - \langle\boldsymbol{y}^\top\hat{\mathbf{\Lambda}}^\top\boldsymbol{M}\Delta\boldsymbol{p}\rangle + \frac{1}{2}\langle\Delta\boldsymbol{p}^\top\boldsymbol{M}\Delta\boldsymbol{p}\rangle\,.\end{aligned}$$

Then, by plugging the definitions of the empirical covariances one is left with

$$\chi^2 = \frac{1}{2}\mathrm{tr}\left[\hat{\mathbf{\Lambda}}^\top\boldsymbol{M}\hat{\mathbf{\Lambda}}\hat{\mathbf{\Omega}}^{\mathrm{d}} - 2\hat{\mathbf{\Lambda}}^\top\boldsymbol{M}\hat{\boldsymbol{R}}^{\mathrm{d}} + \boldsymbol{M}\hat{\mathbf{\Sigma}}\right]. \tag{B.1}$$

B.2. Maximum likelihood estimator.

Proof of Proposition 5.2. In order to derive the MLE, we differentiate (B.1) with respect to $\hat{\boldsymbol{\Lambda}}$:

$$0 = \frac{\partial \chi^2}{\partial \hat{\boldsymbol{\Lambda}}} = \boldsymbol{M}(\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{\Omega}}^{\mathrm{d}} - \hat{\boldsymbol{R}}^{\mathrm{d}}).$$

It is clear that the optimal value of $\hat{\boldsymbol{\Lambda}}_{\mathrm{MLE}}$ does no depend on $\boldsymbol{M}$. To find the explicit expression we equate the derivative to 0 and solve

$$\hat{\boldsymbol{\Lambda}}_{\mathrm{MLE}} = \hat{\boldsymbol{R}}^{\mathrm{d}}(\hat{\boldsymbol{\Omega}}^{\mathrm{d}})^{-1}.$$

To calculate the minimum we can plug this solution into (B.1), obtaining

$$\chi^2_{\mathrm{MLE}} = \frac{1}{2}\mathrm{tr}\left[\boldsymbol{M}\left(\hat{\boldsymbol{\Sigma}} - \hat{\boldsymbol{R}}^{\mathrm{d}}(\hat{\boldsymbol{\Omega}}^{\mathrm{d}})^{-1}(\hat{\boldsymbol{R}}^{\mathrm{d}})^{\top}\right)\right]. \tag{B.2}$$

■

B.3. EigenLiquidity estimator.

The derivation of the EigenLiquidity estimator runs along the same lines as the previous case.

Proof of Proposition 5.3. We want to find an estimator $\hat{\boldsymbol{\Lambda}}_{\mathrm{ELM}}$ that minimizes χ^2 under the constraint $[\hat{\boldsymbol{\Lambda}}_{\mathrm{ELM}}, \hat{\boldsymbol{\Sigma}}] = 0$, which is equivalent to saying that

$$\hat{\boldsymbol{\Lambda}}_{\mathrm{ELM}} = \sum_{a=1}^{n} \hat{\boldsymbol{s}}_a g_a \hat{\boldsymbol{s}}_a^{\top},$$

where $\{\hat{\boldsymbol{s}}_a\}_{a=1}^{n}$ are the eigenvectors of $\hat{\boldsymbol{\Sigma}}$. Because of this constraint, the minimization of χ^2 has to be done with respect to g_a rather than $\hat{\boldsymbol{\Lambda}}$,

$$0 = \frac{\partial \chi^2}{\partial g_a} = \mathrm{tr}\left(\boldsymbol{M}(\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{\Omega}}^{\mathrm{d}} - \hat{\boldsymbol{R}}^{\mathrm{d}})\hat{\boldsymbol{s}}_a\hat{\boldsymbol{s}}_a^{\top}\right),$$

which for arbitrary $\boldsymbol{M}$ implies

$$\hat{\boldsymbol{s}}_a^{\top}(\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{\Omega}}^{\mathrm{d}} - \hat{\boldsymbol{R}}^{\mathrm{d}})\hat{\boldsymbol{s}}_a = 0$$

and therefore

$$g_a = \frac{\hat{\boldsymbol{s}}_a^{\top}\hat{\boldsymbol{R}}^{\mathrm{d}}\hat{\boldsymbol{s}}_a}{\hat{\boldsymbol{s}}_a^{\top}\hat{\boldsymbol{\Omega}}^{\mathrm{d}}\hat{\boldsymbol{s}}_a}.$$

In order to compute the minimum loss we proceed as in the previous case plugging the value of $\hat{\boldsymbol{\Lambda}}_{\mathrm{ELM}}$ into (B.1) and obtain

$$\begin{aligned}
\chi^2_{\mathrm{ELM}} &= \frac{1}{2}\mathrm{tr}\left[M\left(\hat{\boldsymbol{\Sigma}} + (\hat{\boldsymbol{\Lambda}}_{\mathrm{ELM}}\hat{\boldsymbol{\Omega}}^{\mathrm{d}} - 2\hat{\boldsymbol{R}}^{\mathrm{d}})\hat{\boldsymbol{\Lambda}}_{\mathrm{ELM}}\right)\right] \\
&= \frac{1}{2}\mathrm{tr}\left[M\left(\hat{\boldsymbol{\Sigma}} - \hat{\boldsymbol{R}}^{\mathrm{d}}\hat{\boldsymbol{\Lambda}}_{\mathrm{ELM}}\right)\right] \\
&= \frac{1}{2}\mathrm{tr}\left[M\left(\hat{\boldsymbol{\Sigma}} - \hat{\boldsymbol{R}}^{\mathrm{d}}\sum_{a=1}^{n}\hat{\boldsymbol{s}}_a\frac{\hat{\boldsymbol{s}}_a^{\top}\hat{\boldsymbol{R}}^{\mathrm{d}}\hat{\boldsymbol{s}}_a}{\hat{\boldsymbol{s}}_a^{\top}\hat{\boldsymbol{\Omega}}^{\mathrm{d}}\hat{\boldsymbol{s}}_a}\hat{\boldsymbol{s}}_a^{\top}\right)\right].
\end{aligned}$$

■

B.4. Kyle estimator.

Proof of Proposition 5.4. To prove the first statement of the proposition it suffices to compute the covariance of the efficient prices,

$$\hat{\mathbf{\Sigma}}_{\mathrm{Kyle}} = \langle \hat{\Delta} \boldsymbol{p} \hat{\Delta} \boldsymbol{p}^\top \rangle = \hat{\mathbf{\Lambda}}_{\mathrm{Kyle}} \langle \boldsymbol{y}\boldsymbol{y}^\top \rangle \hat{\mathbf{\Lambda}}_{\mathrm{Kyle}}^\top = \hat{\mathbf{\Lambda}}_{\mathrm{Kyle}} \hat{\mathbf{\Omega}}^{\mathrm{d}} \hat{\mathbf{\Lambda}}_{\mathrm{Kyle}}^\top = k^2 \hat{\mathbf{\Sigma}} \,,$$

and to use the unicity result for the above equation (proved in Appendix A.4) in order to recover (5.17).

In order to find the loss at the minimum one has to take into account that the only free parameter is k, and since $\hat{\mathbf{\Lambda}}_{\mathrm{Kyle}} \hat{\mathbf{\Omega}}^{\mathrm{d}} \hat{\mathbf{\Lambda}}_{\mathrm{Kyle}} = k^2 \hat{\mathbf{\Sigma}}$, the loss is

$$\begin{aligned} \chi^2_{\mathrm{Kyle}} &= \frac{1}{2} \mathrm{tr} \left[M \left(\hat{\mathbf{\Sigma}} + (\hat{\mathbf{\Lambda}}_{\mathrm{ELM}} \hat{\mathbf{\Omega}}^{\mathrm{d}} - 2\hat{\boldsymbol{R}}^{\mathrm{d}}) \hat{\mathbf{\Lambda}}_{\mathrm{Kyle}} \right) \right] \\ &= \frac{1}{2} \mathrm{tr} \left[M \left((1+k^2)\hat{\mathbf{\Sigma}} - 2k \hat{\boldsymbol{R}}^{\mathrm{d}} (\hat{\mathcal{R}}^{\mathrm{d}})^{-1} \sqrt{\hat{\mathcal{R}}^{\mathrm{d}} \hat{\mathbf{\Sigma}} \hat{\mathcal{L}}^{\mathrm{d}}} (\hat{\mathcal{L}}^{\mathrm{d}})^{-1} \right) \right] . \end{aligned}$$

Optimizing with respect to k we obtain

$$\frac{\partial \chi^2_{\mathrm{Kyle}}}{\partial k} = k \mathrm{tr} \left[\boldsymbol{M} \hat{\mathbf{\Sigma}} \right] - \mathrm{tr} \left[\boldsymbol{M} \hat{\boldsymbol{R}}^{\mathrm{d}} (\hat{\mathcal{R}}^{\mathrm{d}})^{-1} \sqrt{\hat{\mathcal{R}}^{\mathrm{d}} \hat{\mathbf{\Sigma}} \hat{\mathcal{L}}^{\mathrm{d}}} (\hat{\mathcal{L}}^{\mathrm{d}})^{-1} \right] = 0 \, .$$

■

B.5. Fabricating synthetic covariance matrices inspired from real data. The data set used to fabricate the synthetic covariance matrices in section 5.3 consists of the averaged price variations and volumes in 5 minute bins of the 2-year, 5-year, 10-year, and 30-year tenors of the U.S. Treasury Futures traded in the Chicago Board of Trade during the year 2016.

For each pair of bonds with price variations $\Delta \boldsymbol{p}^{(t)}$ and volumes $\boldsymbol{y}^{(t)}$ we define the normalized covariances as

$$\boldsymbol{C} = \begin{pmatrix} \hat{\mathbf{\Sigma}} & (\hat{\boldsymbol{R}}^{\mathrm{d}})^\top \\ \hat{\boldsymbol{R}}^{\mathrm{d}} & \hat{\mathbf{\Omega}}^{\mathrm{d}} \end{pmatrix} = \boldsymbol{D}^{-1} \begin{pmatrix} \langle \Delta \boldsymbol{p} \Delta \boldsymbol{p}^\top \rangle & \langle \boldsymbol{y} \Delta \boldsymbol{p}^\top \rangle \\ \langle \Delta \boldsymbol{p} \boldsymbol{y}^\top \rangle & \langle \boldsymbol{y}\boldsymbol{y}^\top \rangle \end{pmatrix} \boldsymbol{D}^{-1},$$

where we have previously shifted the data by their empirical means and where

$$\boldsymbol{D} = \mathrm{diag} \left(\sqrt{\langle \Delta p_1^2 \rangle}, \sqrt{\langle \Delta p_2^2 \rangle}, \sqrt{\langle y_1^2 \rangle}, \sqrt{\langle y_2^2 \rangle} \right) .$$

Fixing liquidity. Covariance matrices with the desired liquidity are constructed by rescaling the rows and columns of $\boldsymbol{C}$:

$$\boldsymbol{C}_\epsilon = \begin{pmatrix} \hat{\mathbf{\Sigma}}_\epsilon & (\hat{\boldsymbol{R}}^{\mathrm{d}}_\epsilon)^\top \\ \hat{\boldsymbol{R}}^{\mathrm{d}}_\epsilon & \hat{\mathbf{\Omega}}^{\mathrm{d}}_\epsilon \end{pmatrix} = \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & \sqrt{\epsilon} \end{pmatrix} \boldsymbol{C} \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & \sqrt{\epsilon} \end{pmatrix} .$$

Using this recipe, the variance of the volume of the second asset will be ϵ and consequently the second column of the response is multiplied by $\sqrt{\epsilon}$. Note that the price variations are unchanged.

Fixing correlations. In order to modify the price correlations we have to construct linear combinations of assets. If $\hat{\mathbf{\Sigma}}_{12} = r$, then

$$\boldsymbol{C}_\rho = \begin{pmatrix} \hat{\mathbf{\Sigma}}_\rho & (\hat{\boldsymbol{R}}^{\mathrm{d}}_\rho)^\top \\ \hat{\boldsymbol{R}}^{\mathrm{d}}_\rho & \hat{\mathbf{\Omega}}^{\mathrm{d}}_\rho \end{pmatrix} = \begin{pmatrix} \boldsymbol{A} & 0 \\ 0 & \boldsymbol{A} \end{pmatrix} \boldsymbol{C} \begin{pmatrix} \boldsymbol{A} & 0 \\ 0 & \boldsymbol{A} \end{pmatrix},$$

where

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \sqrt{\frac{1+\rho}{1+r}} & 0, \\ 0 & \sqrt{\frac{1-\rho}{1-r}} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

This modification ensures that $\hat{\mathbf{\Sigma}}_\rho = \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)$. The response and volumes are also modified accordingly. In particular the diagonal entries of $\hat{\mathbf{\Omega}}^{\mathrm{d}}_\rho$ will still be equal (but not necessarily equal to 1).

Note that for $\boldsymbol{C}$ SPD, both $\boldsymbol{C}_\epsilon$ and $\boldsymbol{C}_\rho$ are also SPD and therefore are coherent covariance matrices.

Acknowledgments. We thank Z. Eisler, A. Fosset, and E. Sérié for very fruitful discussions. We also thank J. Caballé for kindly providing the unpublished manuscript [5].

REFERENCES

[1] A. Alfonsi, F. Klöck, and A. Schied, *Multivariate transient price impact and matrix-valued positive definite functions*, Math. Oper. Res., 41 (2016), pp. 914–934.

[2] M. Benzaquen, I. Mastromatteo, Z. Eisler, and J.-P. Bouchaud, *Dissecting cross-impact on stock markets: An empirical analysis*, J. Stat. Mech. Theory Exp., 2017 (2017), 023406.

[3] J.-P. Bouchaud, J. Bonart, J. Donier, and M. Gould, *Trades, Quotes and Prices: Financial Markets Under the Microscope*, Cambridge University Press, Cambridge, 2018.

[4] J. Caballé and M. Krishnan, *Imperfect competition in a multi-security market with risk neutrality*, Econom. J., 62 (1994), pp. 695–704.

[5] J. Caballé and M. Krishnan, *Insider Trading and Asset Pricing in an Imperfectly Competitive Multi-Secrity Market*, Technical report, Unitat de Fonaments de l'Anàlisi Econòmica (UAB) and Institut d'Anàlisi Econòmica (CSIC), 1990.

[6] U. Çetin and A. Danilova, *Markovian Nash equilibrium in financial markets with asymmetric information and related forward-backward systems*, Ann. Appl. Probab., 26 (2016), pp. 1996–2029.

[7] J. Hasbrouck and D. J. Seppi, *Common factors in prices, order flows, and liquidity*, J. Financ. Econ., 59 (2001), pp. 383–411.

[8] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[9] A. S. Kyle, *Continuous auctions and insider trading*, Econom. J., 6 (1985), pp. 1315–1335.

[10] G. Lasserre, *Asymmetric information and imperfect competition in a continuous time multivariate security model*, Finance Stoch., 8 (2004), pp. 285–309.

[11] G. Lindgren, H. Rootzén, and M. Sandsten, *Stationary Stochastic Processes for Scientists and Engineers*, CRC Press, Boca Raton, FL, 2013.

[12] I. Mastromatteo, M. Benzaquen, Z. Eisler, and J.-P. Bouchaud, *Trading lightly: Cross-impact and optimal portfolio execution*, Risk Mag., July (2017), pp. 1–6.

[13] P. Pasquariello and C. Vega, *Strategic cross-trading in the US stock market*, Rev. Finance, 19 (2013), pp. 229–282.

[14] M. Schneider and F. Lillo, *Cross-impact and no-dynamic-arbitrage*, Quant. Finance, 19 (2018), pp. 1–18.

[15] P. Vitale, *Risk-averse insider trading in multi-asset sequential auction markets*, Econom. Lett., 117 (2012), pp. 673–675.

[16] S. Wang, R. Schäfer, and T. Guhr, *Cross-response in correlated financial markets: Individual stocks*, Eur. Phys. J. B, 89 (2016), 105.