March 2014

WORKING PAPER SERIES

2014-ACF-01

Institutional investors and the dependence structure of asset returns

Rama CONT
Imperial College and Université Pierre et Marie Curie

Lakshithe WAGALATH
IESEG School of Management (LEM-CNRS)

IESEG School of Management
Lille Catholic University
3, rue de la Digue
F-59000 Lille
www.ieseg.fr
Tel: 33(0)3 20 54 58 92
Fax: 33(0)3 20 57 48 55

Institutional investors and the dependence structure of asset returns

Rama Cont & Lakshithe Wagalath

Imperial College and Université Pierre et Marie Curie
IESEG School of Management (LEM - CNRS)
DeVinci Finance Lab

February 27, 2014

Abstract

We propose a model of a financial market with multiple assets, which takes into account the impact of a large institutional investor rebalancing its positions, so as to maintain a fixed allocation in each asset. We show that feedback effects can lead to significant excess realized correlation between asset returns and modify the principal component structure of the (realized) correlation matrix of returns. Our study naturally links, in a quantitative manner, the properties of the realized correlation matrix – correlation between assets, eigenvectors and eigenvalues – to the sizes and trading volumes of large institutional investors. In particular, we show that even starting with uncorrelated 'fundamentals', fund rebalancing endogenously generates a correlation matrix of returns with a first eigenvector with positive components, which can be associated to the market, as observed empirically. Finally, we show that feedback effects flatten the differences between assets' expected returns and tend to align them with the returns of the institutional investor's portfolio, making this benchmark fund more difficult to beat, not because of its strategy but precisely because of its size and market impact.

Contents

1 Introduction

International financial markets have become increasingly dominated by large institutional investors, who account for a large fraction of holdings and trades in financial assets. For instance, institutional investors in the US hold $25000 billion in financial assets which represents 17.4% of total outstanding assets. Their positions in the US equity markets amount to $12500 billion, which corresponds to holding 70% of the total equity assets in the US (Gonnard et al., 2008; Tonello and Rabimov, 2010). Two major features characterize large institutional investors over the last years. First, they build their portfolios with the use of indices and exchange-traded funds (representing a sector, a geograpical zone or an asset class for instance), which have become increasingly popular in the last years, and assets traded on large national exchanges (Gastineau, 2010; Fuhr, 2011; Boudreaux, 2012) . Secondly, while such asset managers do not frequently modify their allocations, they do actively trade in the market: Carhart (2012) documents that the average turnover for US mutual funds is 75%.

Large institutional investors build and manage their portfolios comprising numerous assets taking into account the dependence structure between asset returns. In particular, the correlation between asset returns is a key ingredient for trading, portfolio optimization and risk management. It is very often considered as reflecting a structural correlation between fundamentals of asset returns and hence assumed not to vary a lot in time. Ever since Markowitz (1952), theoretical studies show that, under the assumption of a constant correlation structure between asset returns, optimal strategies are fixed-mix strategies, i.e. maintaining a fixed allocation in each asset in the portfolio. Typically, if the value of an asset increased, its weight in the portfolio increases and the investor following a fixed-mix strategy sells part of its positions in this asset, so as to come back to the target allocation for this asset. The fixed-mix strategy implies 'buying low and selling high'. Numerous theoretical studies (Evstigneev and Schenk-Hoppé, 2002; Dempster et al., 2003; Mulvey and Kim, 2008) have shown that such strategies can enhance the long-term growth rate of portfolios.

Whereas the price of financial assets is traditionally modeled as an exogenous stochastic process unaffected by investors' strategies, the presence of institutional investors, typically following fixed-mix strategies and which have a large impact when trading, has implications for financial markets, in particular for the indices and ETFs that they trade, and hence for the components of those indices and ETFs.

Indeed, many empirical studies, such as Aitken (1998); Sias (1996); Sias and Starks (1997), detailed in Section 1.2, document the impact of large institutional investors. They show in particular that trading by institutional investors tends to increase the correlation between the assets that they hold and generate contagion effects to other investors with similar balance sheets. In addition, the study of empirical correlation matrices (Friedman and Weisberg, 1981; Bouchaud et al., 2000) shows that the realized correlation matrix of returns displays common features across stock markets – first eigenvalue which is significantly larger than the others and associated to a 'market' eigenvector, with positive weights on each component, as shown in Figure 1 in the case of the Eurostoxx 50 – which strongly suggest the impact of institutional investors on the dependence structure of asset returns.

In this paper, we develop a quantitative framework for modeling the impact of trading by large institutional investors on the dependence structure of asset returns. Our study shows that rebalancing from institutional investors endogenously increases correlation between asset returns, hence limiting the benefits of diversification and modifying the structure of optimal strategies for investors in this market. Such feedback effects naturally lead to a realized correlation matrix of returns with a first eigenvalue larger than the others and associated to a market eigenvector, as observed empirically. The analytical results that be obtain for realized correlations, asset and fund volatilities, eigenvalues and eigenvectors of the realized correlation matrix, are useful in a risk-management and portfolio allocation perspective, not only for large institutional investors whose impact is modeled in our study, but also for other investors who suffer from contagion effects that we are able to quantify.

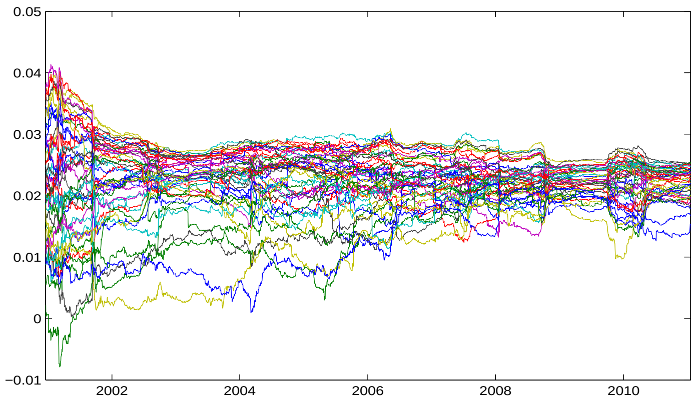

Figure 1: Components of the eigenvector associated to the largest eigenvalue of the empirical correlation matrix of returns for Eurostoxx 50

1.1 Summary

We propose a multi-period model of a financial market with multiple assets, in which a large institutional investor maintains a fixed allocation across assets. Simulations of this model, with realistic parameters estimated from time series of S&P500 stock returns, suggest that feedback effects from the fund's rebalancing lead to a significant increase in realized correlation between asset returns. We exhibit conditions under which the discrete-time model converges to a diffusion limit. By studying the multi-dimensional diffusion limit for the price dynamics, we show that the presence of such large institutional investor maintaining a constant allocation across asset classes may result in a significant and systematic impact on expected returns and the correlation of returns. In particular, such fixed-mix strategies dampen asset volatility but increase correlation across asset classes. This rebalancing effect leads to a systematic bias in the first principal component of the correlation matrix, overweighting assets with high turnover in the benchmark portfolio. In particular, in the presence of feedback effects, and even starting from uncorrelated fundamentals, trading by the large fund endogenously generates a realized correlation matrix of returns with a first eigenvector with positive components. The impact of the large institutional investor biases asset expected returns and decreases the performance of funds who overweigh (resp. underweigh) assets with large (resp. low) expected returns. These findings have consequences for risk-management and asset allocation. We show that the impact of the large institutional investor modifies the risk/return trade-off of portfolios composed from the same assets: an investor who factors these effects into his allocations can improve his risk/return trade-off.

1.2 Related literature

Various empirical studies attest to the large market share of institutional investors. Gonnard et al. (2008) study institutional investors of countries of the Organization for Economic Co-operation and Development (OECD) while Tonello and Rabimov (2010) focus on the institutional investors in the US. Such investors comprise mutual funds, insurance companies and pension funds. Their investments amounted to \$40000 billion for funds in the OECD in 2005, which represents 150% of the gross domestic product of the OECD. US institutional investors account for more than half of those investments (\$25000 billion) and prefer investing in equity markets (50% of their positions).

The preferences of large institutional investors are examined in numerous empirical studies. Del Guercio (1996) finds empirically that banks, contrary to mutual funds, prefer investing in prudent stocks. Gompers and Metrick (2001) use a database with seventeen years of data on large institutional investors and show that they prefer holding liquid assets, while Ferreira and Matos (2008) find that institutional investors have a strong preference for the stocks of large firms and firms with good governance. Falkenstein (2012) shows that mutual funds prefer investing in liquid stocks with low transaction costs and are averse to stocks with low idiosyncratic volatility. Lakonishok et al. (1992) study the types of strategies followed by institutional investors and whether they follow trading practices which are potentially destabilizing for asset prices. The impact of institutional investors on asset returns is widely studied in the empirical literature. Aitken (1998) shows that the growth of capital invested by mutual funds and other institutional investors in emerging markets resulted in a sharp increase of autocorrelation for the assets in those markets. Sias and Starks (1997) also finds that the larger the institutional ownership of a stock in the NYSE, the larger its autocorrelations while Sias (1996) finds that an increase in ownership by institutional investors on a given stock results in a greater stock volatility.

Most theoretical studies model the impact of large institutional investors on a single asset's return and volatility. Almgren and Chriss (2000); Almgren and Lorenz (2006); Almgren (2009) model the permanent and temporary impact of a large investor liquidating a position on a single asset and derive an optimal liquidation strategy. Alfonsi et al. (2009) derive the optimal strategy to liquidate a large position on an asset by taking into account the order book of the asset. Gabaix et al. (2006) propose an equilibrium model which takes into account the supply and demand of a large institutional investor and show that their trades can lead to volatility spikes. They derive an optimal strategy for the institutional investor, in the presence of its own feedback effects. These theoretical studies explain quantitatively the facts described in the empirical literature cited previously, but mainly focus on a single asset and derive optimal strategies for institutional investors. In particular, they do not model the cross-asset impact of large institutional investors and the spillover effects that they can generate. Kyle and Xiong (2001) study a market with two risky assets and three types of traders: noise traders, convergence traders and long-term investors. They show how the strategies implemented by each type of traders can result in contagion effects and lead to endogenous correlation which can not be explained by assets fundamentals.

Our quantitative results show that, even starting from homoscedastic fundamentals, rebalancing by institutional investors naturally generate heteroscedasticity in the covariance structure of asset returns, hence giving an economic explanation for the variability of correlations and volatilities observed in financial markets and which have been modeled, in an exogenous way, by various theoretical studies (Engle, 2002; Da Fonseca et al., 2008; Gouriéroux et al., 2009; Stelzer, 2010; Da Fonseca et al., 2013). In addition, we find that in the presence of feedback effects from institutional investors, the realized correlation matrix of returns displays the features verified by empirical correlation matrices: the first eigenvalue is significantly higher than the other eigenvalues and is associated to an eigenvector with positive components. While studies using random matrix theory such as Bouchaud et al. (2000) are able to explain the emergence of a dominant eigenvalue for realized correlation matrices, such studies generally lead to a first eigenvector which is invariant by rotation, and hence fail to account for the fact that the first eigenvector is a 'market' eigenvector, with positive weights on each component.

1.3 Outline

This paper is organized as follows: Section 2 presents a framework for modeling the impact of trading by a large institutional investor on asset returns. Section 3 studies the dependence structure of asset returns in the presence of feedback effects from the large institutional investor. Section 4 analyzes the impact of the large institutional investor on the risk and returns of assets and for other investors in the market.

2 Asset dynamics in the presence of a large institutional investor

2.1 Multi-period model

Asset fundamentals
Consider a discrete-time market, where trading takes place at dates $t_k = k\Delta t$ and which comprises n financial assets. The value of asset i at t_k is S^i_k. Typically, one can consider that S^i is the value of an index or an ETF representing a sector, asset class or geographic zone. Between t_k and t_{k+1}, the value of each asset moves due to 'fundamentals',

represented by an IID sequence $(\xi_{k+1})_{k\geq 0} = (\xi^1_{k+1}, ..., \xi^n_{k+1})_{k\geq 0}$ of centered random variables with covariance matrix Σ. In the absence of other effects, the log-return of asset i between t_k and t_{k+1} would be

$$\Delta t(m_i - \frac{\Sigma_{i,i}}{2}) + \sqrt{\Delta}t\xi^i_{k+1}$$

where m_i is the 'fundamental' expected return of asset i and Σ is the 'fundamental' covariance matrix of asset returns. Σ reflects the fundamental structure between the n assets in the market. While most practitioners consider that the covariance matrix is constant over a time horizon typically of one year, more statistically-sophisticated models have been proposed for modeling heteroscedasticity and the variability of the fundamental covariance matrix (Engle, 2002; Da Fonseca et al., 2008; Gouriéroux et al., 2009; Stelzer, 2010; Da Fonseca et al., 2013). In this paper, for clarity purpose only, we assume that the fundamental covariance matrix Σ is contant. Our results can be extended to the case of a stochastic fundamental covariance matrix by means of more technicalities, which do not bring more intuition to the problem raised by this paper, namely disentangling the impact of fundamentals and fund rebalancing on the dependence structure of asset returns.

The large institutional investor
Consider now a large institutional investor/fund investing in this market and following a (long) *fixed-mix strategy*, maintaining a fixed allocation in each asset. As discussed in Section 1, the fixed-mix strategy is widely used by institutional investors, in between two allocation dates, which correspond to a time frame of several months. At each date t_k, *the fund holds a (constant, positive) proportion* x_i *of each asset* i which means that the dollar amount invested by the fund in asset i at this date is equal to $x_i W_k$ where W_k is the fund value/wealth at t_k. Denoting by ϕ^i_k the number of units of asset i held at t_k, the fixed-mix strategy implies that at each date:

$$\phi^i_k S^i_k = x_i W_k \tag{1}$$

At each period, the asset values may change due to fundamentals and the fund rebalances its positions in order to maintain the target proportion x_i in each asset i. If the value of an asset increased (resp. decreased) more than the others, the fund sells (resp. buys) units of this asset in order to maintain a fixed portion of this asset in its portfolio. The fixed-mix strategy is a typical example of a contrarian strategy, which implies 'buying low and selling high'. The rebalancing by the fund, in order to maintain its target allocation, generates a net demand of $\phi^i_{k+1} - \phi^i_k$ units of asset i between t_k and t_{k+1}, in a self-financing manner:

$$W_{k+1} = \sum_{i=1}^{n} \phi^i_{k+1} S^i_{k+1} = \sum_{i=1}^{n} \phi^i_k S^i_{k+1} \tag{2}$$

Price impact
The rebalancing of large positions by the fund – this is the case when the fund is large –

impacts prices in a non-random manner. The impact of large orders on asset returns has been modeled in numerous manners: linear (Obizhaeva, 2011; Cont et al., 2013), square root (Bence et al., 2011), or more generally concave (Almgren et al., 2005; Moro et al., 2009). We assume that the impact of this net demand by the fund on the return of each asset i is linear and is measured by the depth D_i of the market in asset i: a net demand of $\frac{D_i}{100}$ shares for security i moves the price of i by one percent. Note that we choose a linear price impact for clarity purpose only, as a general price impact function leads to the same continuous-time limit as a linear price impact, as shown in (Cont and Wagalath, 2013a).

The discrete-time price dynamics
The value of asset i at date t_{k+1} has to verify:

$$S^i_{k+1} = S^i_k \underbrace{\exp\left(\Delta t(m_i - \frac{\Sigma_{i,i}}{2}) + \sqrt{\Delta} t \xi^i_{k+1}\right)}_{'fundamentals'} \times \underbrace{\left(1 + \frac{\phi^i_{k+1} - \phi^i_k}{D_i}\right)}_{feedback\ from\ the\ large\ investor} \quad (3)$$

However, as ϕ^i_{k+1} depends on S^i_{k+1} given Eq (1), we have to prove that, at each period, the fund can rebalance its positions in a self-financing way so as to keep its fixed allocation ie: verify Eq (3). This is done in the following Proposition.

Proposition 2.1 *There exists a unique investment strategy which enables the fund to keep a constant proportion x_i invested in each asset i at every period k, in a self-financing way and (S, W) verify Equations 1, 2 and 3.*

The proof of this proposition is given in Appendix 5.1. At each period, the return of an asset can be decomposed into a fundamental component and a systematic component which is generated endogenously by the large fund's rebalancing. We remark that when market depths are infinite ($D_i = \infty$), the fund's rebalancing does not generate any feedback on asset returns and asset values move according to 'fundamentals' only, which are captured by the fundamental covariance matrix Σ and the fundamental expected returns m.

2.2 Simulation experiments

In this section, we present some results of simulation experiments which illustrate the impact of feedback effects from a large institutional investor following a fixed-mix strategy on the realized correlation between asset returns and the principal component properties of the realized correlation matrix of returns.

Choice of parameters We simulate the multi-period model in a very simple example of homogenous fundamental volatility, correlation and expected return. In order to compare our numerical results with empirical results on the S&P500 in 2006, we choose the following realistic parameters. The simulated market comprises $n = 500$ assets and trading is possible everyday ($\Delta t = \frac{1}{250}$). Each asset has a fundamental expected return $m_i = 11\%$ (equal to the return of the S&P500 in 2006) and a fundamental volatility $\sqrt{\Sigma_{i,i}} = 10\%$ (equal to the realized volatility of the S&P500 in 2006). We denote by $\rho = \frac{\Sigma_{i,j}}{\sqrt{\Sigma_{i,i}\Sigma_{j,j}}}$ the fundamental correlation between any pair of assets.

We consider an institutional investor maintaining a constant portion $x_i = \frac{1}{n} = 0.2\%$ invested in each asset i and the initial position of the fund in each asset is equal to one fifth of the respective asset market depth: for all i, $\frac{\phi_0^i}{D_i} = \frac{1}{5}$. This choice is legitimated by empirical studies: Tonello and Rabimov (2010) shows that the size of institutional investors over the last years is approximately \$25000 billion, among which 50%, ie: \$12500 billion are invested in US equity markets. As a proxy of market depth, we use, following Obizhaeva (2011),

$$D = \frac{Average\ Daily\ Volume}{0.33 \times Daily\ Volatility}$$

Given that in 2006 the average daily volume of the US equity market was \$80 billion and the realized volatility of the S&P500 was 10%, we find that the depth of the US equity markets is $\frac{80}{0.33 \times \frac{10\%}{\sqrt{250}}} \approx \$38000$ billions. Assuming, for example, that 60% of the institutional investors follow a fixed-mix strategy, this legitimates our choice of $\frac{\phi_0^i}{D_i} = 60\% \times \frac{12500}{38000} = \frac{1}{5}$.

Note that we choose to compare our numerical results to empirical results on the S&P500 in 2006 because the recent financial turmoil started in 2007 and, since then, the US equity market was subject to major fire sales and deleveraging phenomenon, which were the main source of feedback in those markets, as analysed in Cont and Wagalath (2013a).

For each of the 10000 simulated scenarios, we compute the realized volatility of asset returns and the realized correlation between asset returns. The following figures display the distribution of those quantities, in the case of $\rho = 0$. We compare those distributions to the case without feedback effects, which corresponds to the case when market depths are infinite.

Realized correlation and realized variance Figure 2 shows that the institutional investor's strategy increases realized correlation between asset returns. Whereas without feedback effects, the distribution of realized correlation between the two assets is centered around its fundamental value $\rho = 0$, we witness, in the presence of feedback effects, that the distribution of realized correlation is shifted towards positive values, centered around an average value of 10% and with values over 20% with significant probability. Even starting with zero fundamental correlations, feedback effects generate, on average, a realized correlation of 10% between assets. In addition, numerical experiments show that the price impact of fixed-mix strategies decreases asset volatility. This is due to the fact that the fund follows a 'contrarian' strategy: it buys (resp. sells) assets that decreased (resp. increased) the most, hence dampening their decrease (resp. increase) and, overall, dampening the amplitude of price moves.

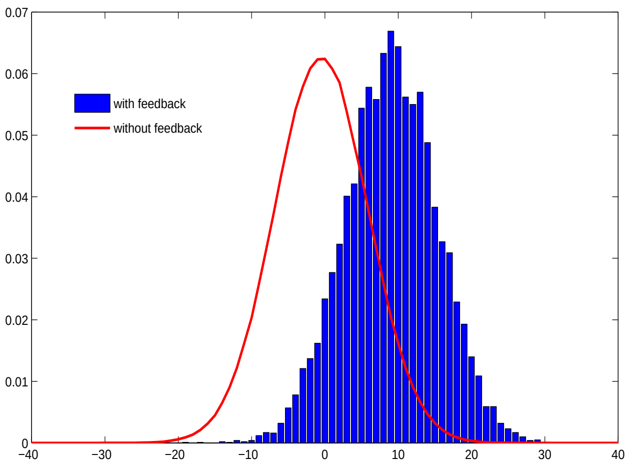

Figure 2: Distribution of realized correlation between assets 1 and 2 (with $\rho = 0$) with and without feedback effects

Average pairwise realized correlation and highest eigenvalue of the realized correlation matrix Table 1 displays the average for the average pairwise realized correlation over 10^4 simulations, for different values of fundamental correlation ρ and fund sizes $\frac{\phi_0^i}{D_i}$. It shows that, for each choice of parameters, due to the impact of the large investor, the average pairwise correlation is higher than its fundamental value in the presence of feedback effects. Furthermore, we see that the larger the fund's positions as a fraction of asset depth (ie: the larger the fund's positions or the lower the assets' depth or liquidity), the larger the impact on the average pairwise correlation. In 2006, the average pairwise one-year realized correlation in the S&P500 was 21%. Table 1 shows that *with a reasonable and realistic choice of parameters for the fund's size*, for example $\frac{\phi_0^i}{D_i} = \frac{1}{5}$ as discussed in the beginning of this section, *an homogenous fundamental correlation of only 15% combined to feedback effects generated by the rebalancing of the fund's positions generate the 22% average pairwise realized correlation observed empirically.*

Table 2 leads to the same conclusions: the presence of the large institutional investor increases the value of the largest eigenvalue of the realized correlation matrix, compared to its fundamental value which is given in the column $\frac{\phi_0^i}{D_i} = 0$. Furthermore, the larger the fund's positions as a fraction of market depth, the larger the eigenvalue of the realized correlation matrix. While the largest eigenvalue of the one-year realized correlation matrix of the S&P500 in 2006 was 110, we see that our model leads to this level of largest eigenvalue starting from a fundamental correlation of 15%, which corresponds to a fundamental largest eigenvalue of 76, combined to feedback effects, in the case where $\frac{\phi_0^i}{D_i} = \frac{1}{5}$.

Average pairwise realized correlation			
ρ	$\frac{\phi_0^i}{D_i} = \frac{1}{10}$	$\frac{\phi_0^i}{D_i} = \frac{1}{5}$	$\frac{\phi_0^i}{D_i} = \frac{1}{3}$
0	2%	5%	10%
10%	12%	15%	21%
15%	18%	22%	29%
25%	29%	35%	44%
50%	55%	61%	69%
75%	79%	82%	87%
90%	92%	93%	95%

Table 1: Average for the average pairwise realized correlation for different values of fundamental correlation ρ and fund sizes (as a fraction of market depth) $\frac{\phi_0^i}{D_i}$. In comparison, the average pairwise one-year realized correlation of the S&P500 in 2006 was 21%.

Largest eigenvalue of the realized correlation matrix				
ρ	$\frac{\phi_0^i}{D_i} = 0$	$\frac{\phi_0^i}{D_i} = \frac{1}{10}$	$\frac{\phi_0^i}{D_i} = \frac{1}{5}$	$\frac{\phi_0^i}{D_i} = \frac{1}{3}$
0	1	8	9	11
10%	51	64	79	105
15%	76	93	113	146
25%	126	149	175	216
50%	251	277	307	345
75%	375	394	412	434
90%	450	459	467	476

Table 2: Average for the largest eigenvalue of the realized correlation matrix for different values of fundamental correlation ρ and fund sizes (as a fraction of market depth) $\frac{\phi_0^i}{D_i}$. In comparison, the largest eigenvalue of the one-year realized correlation matrix of the S&P500 in 2006 was 110.

2.3 Continuous-time limit

We now analyze the continuous-time limit of the multi-period model: the study of this limit enables to obtain analytical formulas for realized correlation between asset returns, eigenvalues and eigenvectors of the realized correlation matrix and asset expected returns, which confirm quantitatively the effects observed in the numerical experiments.

The following theorem describes the diffusion limit of the price process.

Theorem 2.2 *Under Assumption 5.2 given in Appendix 5.2,* $(S_{\lfloor \frac{t}{\Delta t} \rfloor}, W_{\lfloor \frac{t}{\Delta t} \rfloor})_{t \geq 0}$ *converges weakly to a diffusion process* $(P_t, V_t)_{t\geq 0} = (P_t^1, ..., P_t^n, V_t)_{t\geq 0}$ *as* Δt *goes to 0 with:*

$$dP_t^i = b_i(P_t, V_t)dt + (a(P_t, V_t)dB_t)_i \qquad 1 \leq i \leq n$$
$$dV_t = b_{n+1}(P_t, V_t)dt + (a(P_t, V_t)dB_t)_{n+1}$$

where a *and* b *are defined in Appendix 5.2 in Eq* (27) *and Eq* (28) *respectively and* B_t *is an* n*-dimensional Brownian motion.*

In the continuous-time limit, at each date t, the fund allocates x_i to asset i. Its holdings in asset i are:

$$\phi_t^i = \frac{x_i V_t}{P_t^i} \tag{4}$$

The proof of this theorem is given in Appendix 5.2. a and b can be computed explicitely from Lemma 5.4 and Lemma 5.5 in Appendix 5.2. Theorem 2.2 enables to quantify, in a tractable manner, the impact of the large fund on asset dynamics. In the next sections, in order to characterize this impact more intuitively, we will study the impact at order one in liquidity, which enables us to decompose realized correlations, eigenvalues, eigenvectors and fund volatility into a stationary fundamental part and a liquidity-dependent and path-dependent part, generated endogenously by the large institutional investor.

In the case where market depths are infinite (for all i, $D_i = +\infty$), the expression for a and b simplifies to

$$a_{i,k}(P, V) = P^i A_{i,k} \quad and \quad b^i(P, V) = P^i m_i \qquad 1 \leq i \leq n$$

where A is a square root of Σ and the price follows a multivariate Black and Scholes dynamics with expected return m and covariance matrix Σ:

$$\frac{P_t^i}{P_0^i} = \exp((m_i - \frac{\Sigma_{i,i}}{2})t + (AB_t)_i) \quad and \quad \frac{V_t}{V_0} = \exp((X.m - \frac{X.\Sigma X}{2})t + X.AB_t) \tag{5}$$

When market depths are finite, feedback effects from the large fund modify this fundamental price dynamics, as described in the next sections.

3 Dependence structure of asset returns

In this section, we study the impact of the large fund's rebalancing on the dependence structure of asset returns and we show that this dependence structure may depend more on the type and market cap of the various strategies used by large market players than on fundamentals. The fund's impact can be quantified by studying the realized covariance/correlation matrix of asset returns (Andersen et al., 2003; Barndorff-Nielsen and Shephard, 2004).

3.1 Realized covariance and correlation between asset returns

Proposition 3.1 shows that realized covariances and correlations can be decomposed into a fundamental component and an endogenous component generated by the fund's rebalancing strategy. When the rebalancing volumes are large compared to the depth of assets, this endogenous component is exacerbated and can become significant compared to the fundamental component.

The realized covariance matrix of asset returns between t_1 and t_2, denoted $C_{[t_1,t_2]}$, is defined by

$$C^{i,j}_{[t_1,t_2]} = \frac{1}{t_2 - t_1}\left([\ln P^i, \ln P^j]_{t_2} - [\ln P^i, \ln P^j]_{t_1}\right)$$

where $[\ln P^i, \ln P^j]_t$ is the quadratic covariation between $\ln P^i$ and $\ln P^j$ on $[0,t]$ and the realized correlation between t_1 and t_2 is defined by

$$R^{i,j}_{[t_1,t_2]} = \frac{C^{i,j}_{[t_1,t_2]}}{\left(C^{i,i}_{[t_1,t_2]} C^{j,j}_{[t_1,t_2]}\right)^{\frac{1}{2}}}$$

Proposition 3.1 *The realized covariance and realized correlation between assets i and j returns can be respectively decomposed as follows:*

$$C^{i,j}_{[0,T]} = \Sigma_{i,j} + \Lambda_i \Delta\Phi^i_T \sum_{l=1}^{n} x_l(\Sigma_{j,l} - \Sigma_{i,j}) + \Lambda_j \Delta\Phi^j_T \sum_{l=1}^{n} x_l(\Sigma_{i,l} - \Sigma_{i,j}) + O(\|\Lambda\|^2) \quad (6)$$

$$R^{i,j}_{[0,T]} = \frac{\Sigma_{i,j}}{\sqrt{\Sigma_{i,i}\Sigma_{j,j}}} + \frac{\Lambda_i \Delta\Phi^i_T}{\sqrt{\Sigma_{i,i}\Sigma_{j,j}}} \sum_{l=1}^{n} x_l\left(\Sigma_{j,l} - \frac{\Sigma_{i,j}}{\Sigma_{i,i}}\Sigma_{i,l}\right) + \frac{\Lambda_j \Delta\Phi^j_T}{\sqrt{\Sigma_{i,i}\Sigma_{j,j}}} \sum_{l=1}^{n} x_l\left(\Sigma_{i,l} - \frac{\Sigma_{i,j}}{\Sigma_{j,j}}\Sigma_{j,l}\right) + O(\|\Lambda\|^2) \quad (7)$$

where Λ is an n dimensional vector representing the initial holdings of the fund in asset i as a fraction of market depth:

$$\Lambda_i = \frac{\phi^i_0}{D_i} \quad (8)$$

and $\Delta\Phi^i_T$ is a measure of the volume traded on asset i by the large fund:

$$\Delta\Phi^i_T = 1 + \int_0^T \left(1 - \frac{s}{T}\right) \frac{d\phi^i_s}{\phi^i_0} = \frac{1}{T}\int_0^T \frac{\phi^i_s}{\phi^i_0} ds > 0 \quad (9)$$

and $\mathbb{E}\left(\frac{O(\|\Lambda\|^2)}{\|\Lambda\|^2}\right)$ *is bounded when* Λ *goes to zero.*

The proof of this corollary is given in Appendix 5.3. In the presence of feedback effects, the realized covariance/correlation matrix is the sum of the fundamental covariance/correlation matrix and an excess realized covariance/correlation matrix generated by the impact of the fund's rebalancing strategy.

The magnitude of the institutional investor's impact on asset returns is naturally measured by the quantities

$$\Lambda_i = \frac{\phi_0^i}{D_i}$$

which measure the size of the institutional investor's positions in each asset, as a fraction of asset market depth and Proposition 3.1 gives the expansion at order one in $\|\Lambda\|$ of the realized covariances and correlations.

Equation 7 shows that realized correlations between asset returns are impacted by the strategy followed by the institutional investor, namely

- its size Λ
- its allocations x_i
- its rebalancing strategy $\Delta\Phi^i$

They can deviate significantly from fundamentals when the size of the fund's positions and trading volumes are large compared to asset market depths. As the size of institutional investors and their market shares has become very large, as discussed in Section 1, Proposition 3.1 shows that *the dependence structure of asset returns may depend more on the type and market cap of the various strategies used by large market players than on fundamentals.*

We remark that when the fund invests significantly in one asset i_0, such that $\Lambda_{i_0} > 0$, and even if its positions on the other assets are negligible ($\Lambda_i = 0$ for $i \neq i_0$), the realized covariance between those assets and i_0 is modified and different from the fundamental covariance, generating a contagion effect from this asset to all other assets held by the fund.

3.2 Case of zero fundamental correlations

In this section, we examine the case when the fund invests in assets with zero fundamental correlations: the fundamental covariance matrix Σ is diagonal and we write:

$$\Sigma_{i,i} = \sigma_i^2 \quad and \quad \Sigma_{i,j} = 0 \quad i \neq j$$

We show that rebalancing by the large fund endogenously generates *positive realized correlation between assets with zero fundamental correlation*, reducing the benefits of diversification for the fund. The presence of the fund induces a *minimum structural realized correlation* between asset returns, which depends on liquidity and rebalancing flows. Even starting from a fundamental correlation matrix which is the identity, feedback effects from the large fund naturally lead to the empirically-observed structure for realized correlation matrices: a *dominant eigenvalue, which is of the order of the number of assets in the market n*, and associated to an *eigenvector with strictly positive components, corresponding to the market.*

Using Eq (7), we can compute the realized correlation between assets i and j returns:

$$R^{i,j}_{[0,T]} = \Lambda_i \frac{\sigma_j x_j}{\sigma_i} \Delta\Phi^i_T + \Lambda_j \frac{\sigma_i x_i}{\sigma_j} \Delta\Phi^j_T + O(\|\Lambda\|^2) > 0 \tag{10}$$

Eq (9) implies that $\Delta\Phi^i_T > 0$ and we deduce that feedback effects from the fund's rebalancing generates positive realized correlation between assets with zero fundamental correlations. Intuitively, this stems from the fact that the systematic strategy used by the fund creates a similar pattern of behavior for all assets and shows that diversification effects are reduced by the fund's own impact. This analytical result confirms quantitatively the numerical results of Figure 2. Excess correlation is exacerbated by the size of the fund relative to market depth ($\|\Lambda\|$) and the volumes of rebalancing ($\Delta\Phi$), which are endogenously determined by the fund's strategy $(x_1, ..., x_n)$. Our results show that *when liquidity effects are accounted for, observed levels of realized correlation across assets are compatible with the null hypothesis of absence of correlation in fundamentals.*

Eq (10) enables us to derive a structural minimum for the realized correlation between a given pair of assets i and j:

$$R^{i,j}_{[0,T]} \geq 2\sqrt{x_i \Lambda_i x_j \Lambda_j} (\Delta\Phi^i_T \Delta\Phi^j_T)^{\frac{1}{2}} + O(\|\Lambda\|^2)$$

This lower bound for realized correlation is triggered by the mere presence of the institutional investor following a fixed-mix strategy. In particular, the more the fund invests in two assets, the greater the correlation it generates endogenously between the two assets.

As the realized correlation matrix $R_{[0,T]}$ has strictly positive terms, the Perron Frobenius theorem states that it has an eigenvalue which is strictly higher than its other eigenvalues and which is associated to an eigenvector with strictly positive coordinates. Furthermore, this eigenvalue belongs to the interval:

$$\left[1 + (n-1) \underbrace{\left(\frac{1}{n(n-1)} \sum_{i \neq j} R^{i,j}_{[0,T]} \right)}_{\text{average pairwise correlation}} \; ; \; 1 + \max_i \sum_{j \neq i} R^{i,j}_{[0,T]} \right] \tag{11}$$

What is remarkable is that, even starting with a fundamental correlation matrix which is equal to the identity, and hence has all eigenvalues equal to 1 and no particular structure for eigenvectors, the impact of the fund naturally generates a correlation matrix with a dominant eigenvalue, of the order of n and associated to a 'market' eigenvector. Our results enable to explain quantitatively the existence of such a market vector associated to the largest eigenvalue of the correlation matrix, as illustrated in Figure 1, while, as discussed in the literature review of Section 1, existing studies modeling the principal component properties of the realized correlation matrix of returns fail to explain this empirically-observed feature.

Finally, using Corollary 3.1, we find that the realized variance of asset i returns is given by

$$C^{i,i}_{[0,T]} = \sigma_i^2 \left(1 - 2(1-x_i)\Delta\Phi^i_T\right) + O(\|\Lambda\|^2) < \sigma_i^2$$

which implies that feedback effects decrease the realized variance of asset returns. This is consistent with the fact that the fund buys (resp. sells) assets which decreased (resp. increased) the most, dampening the amplitude of asset movements.

3.3 Eigenvalues and eigenvectors of the realized correlation matrix

We now analyze the principal component properties of the realized correlation matrix of returns, in the case of a general fundamental correlation matrix. Our study enables to establish a quantitative link between the eigenvalues and eigenvectors of the correlation matrix and the strategy implemented by the fund.

As the realized correlation matrix in the presence of feedback effects can be considered as a perturbation of the fundamental correlation matrix, one can expect that its eigenvalues and eigenvectors are also perturbations of the corresponding eigenvalues and eigenvectors of the fundamental correlation matrix. Recall that the realized correlation matrix $R_{[0,T]}$ and the fundamental correlation matrix $\overline{R}$ are real-valued matrices defined respectively by:

$$R^{i,j}_{[0,T]} = \frac{C^{i,j}_{[0,T]}}{\left(C^{i,i}_{[0,T]} C^{j,j}_{[0,T]}\right)^{\frac{1}{2}}} \quad and \quad \overline{R}^{i,j} = \frac{\Sigma^{i,j}}{\left(\Sigma^{i,i}\Sigma^{j,j}\right)^{\frac{1}{2}}} \tag{12}$$

As they are both symmetric and non-negative, we know that there exist eigenvalues $v_1 \geq ... \geq v_n \geq 0$, $\overline{v}_1 \geq ... \geq \overline{v}_n \geq 0$ and two orthonormal bases $(\psi_1, ..., \psi_n)$, $(\overline{\psi}_1, ..., \overline{\psi}_n)$ such that for all $1 \leq j \leq n$:

$$R_{[0,T]}\psi_j = v_j \psi_j \quad and \quad \overline{R\psi}_j = \overline{v}_j \overline{\psi}_j \tag{13}$$

The following proposition gives analytical formulas which quantify the impact of the large institutional investor on the principal component properties of the realized correlation matrix $R_{[0,T]}$.

Assumption 3.2 *$\overline{v}_j$ defined in Eq (13) is a simple eigenvalue for the fundamental correlation matrix $\overline{R}$.*

Proposition 3.3 *Under Assumption 3.2, there exists $\gamma > 0$ such that if $\|\Lambda\| < \gamma$, then v_j is a simple eigenvalue for the realized correlation matrix $R_{[0,T]}$ and:*

$$v_j = \overline{v_j} +^t \overline{\psi_j}\ \nabla R\ \overline{\psi_j} + o(\|\Lambda\|)$$

and v_j is associated to the unit eigenvector ψ_j, which is collinear to:

$$\overline{\psi}_j + \sum_{k \neq j} \frac{{}^t\overline{\psi_k}\ \nabla R\ \overline{\psi_j}}{\overline{v_j} - \overline{v_k}} \overline{\psi_k} + o(\|\Lambda\|)$$

where v_j, $\overline{v}_j$, ψ_j, $\overline{\psi}_j$ are defined in Eq (13)*, ∇R is a symmetric matrix defined by $[\nabla R]_{i,i} = 0$ and for $i \neq j$:*

$$[\nabla R]_{i,j} = \frac{\Lambda_i \Delta\Phi_T^i}{\sqrt{\Sigma_{i,i}\Sigma_{j,j}}} \sum_{l=1}^{n} x_l(\Sigma_{j,l} - \frac{\Sigma_{i,j}}{\Sigma_{i,i}}\Sigma_{i,l}) + \frac{\Lambda_j \Delta\Phi_T^j}{\sqrt{\Sigma_{i,i}\Sigma_{j,j}}} \sum_{l=1}^{n} x_l(\Sigma_{i,l} - \frac{\Sigma_{i,j}}{\Sigma_{j,j}}\Sigma_{j,l})$$

and $\frac{o(\|\Lambda\|)}{\|\Lambda\|}$ converges almost surely to zero when Λ goes to zero.

Proposition 3.3 is a well-know result of perturbation theory for eigenvalues and eigenvectors of symmetric matrices, analogous to the result of (Allez and Bouchaud, 2012, Eq 2.2). It gives a tractable formula for the eigenvalues and eigenvectors of the realized correlation matrix in the presence of feedback effects from the investor following the fixed-mix strategy $(x_1, ..., x_n)$. Due to such feedback effects, the eigenvalues and eigenvectors of the realized correlation matrix depend on the sizes and allocations of institutional investors.

This impact naturally modifies the profile of risk in the market for all investors, and for the large institutional investor itself. Assume for example that the institutional investor, in order to reduce its volatility, has chosen a strategy $(x_1, ..., x_n)$ so as to stay orthogonal to $\overline{\psi_1}$, the first eigenvector of the fundamental correlation matrix, associated to the largest eigenvalue and hence the largest regime of volatility. As a consequence, $(x_1, ..., x_n)$ is chosen in the subspace generated by $\overline{\psi_2}, ..., \overline{\psi_n}$. However, the impact of the fund's rebalancing generates a change in the first eigenvalue, equal to $\sum_{k=2}^{n} \frac{{}^t\overline{\psi_k}\ \nabla R\ \overline{\psi_1}}{\overline{v_1} - \overline{v_k}} \overline{\psi_k}$ whose direction is precisely in the subspace generated by $\overline{\psi_2}, ..., \overline{\psi_n}$, containing the vector of allocations $(x_1, ..., x_n)$, hence generating larger-than-expected realized volatility.

Example and numerical tests In this paragraph, we illustrate how feedback effects impact the principal component properties of the realized correlation matrix of returns in a simple example where all parameters are homogenous: asset fundamental volatilities are equal to σ, fundamental correlation between any pair of assets is ρ, the large institutional investor's allocation in each asset i is $x_i = \frac{1}{n}$ and the size of its position in each asset as a fraction of market depth is $\Lambda_i = \overline{\Lambda}$.

Corollary 3.4 *Under the assumption of homogenous fundamental asset volatilities σ, correlations ρ, allocations $x_i = \frac{1}{n}$ and ratios holdings to market depths $\Lambda_i = \Lambda$, the largest eigenvalue of the realized correlation matrix of returns is equal to:*

$$v_1 = 1 + (n-1)\rho + 2\Lambda(1-\rho)\left(\rho + \frac{1-2\rho}{n} - \frac{1-\rho}{n^2}\right)\sum_{j=1}^{n} \Delta\Phi_T^j + o(\Lambda) \qquad (14)$$

and is associated to the first eigenvector ψ_1 *which is proportional to:*

$$\alpha_n(x,\Lambda)\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} + \beta n(x,\Lambda)\begin{pmatrix}\Delta\Phi_T^1\\ \vdots\\ \Delta\Phi_T^i\\ \vdots\\ \Delta\Phi_T^n\end{pmatrix} + o(\Lambda) \tag{15}$$

where

$$\alpha_n(x,\Lambda) = 1 + \overline{\Lambda}(1-\rho)(-\rho + \frac{3\rho-1}{n} + \frac{2(1-\rho)}{n^2})\sum_{j=1}^{n}\Delta\Phi_T^j$$

$$\beta_n(x,\Lambda) = (n-2)\overline{\Lambda}(1-\rho)(\rho + \frac{1-\rho}{n})$$

Eq 14 shows that feedback effects increase the value of the largest eigenvalue of the realized correlation matrix. We simulated 10^4 price trajectories of our model and, for each trajectory, we calculated numerically the largest eigenvalue of the realized correlation matrix and computed the theoretical largest eigenvalue given by Eq (14). Table 3 shows that the average error made by using the theoretical formula of Eq (14) to estimate the largest eigenvalue of the realized correlation matrix is 10%, which is significantly lower than when using the fundamental largest eigenvalue (error of 60%, using fundamental largest eigenvalue, which is equal to $1+(n-1)\rho$).

Eq 15 gives the structure of the first eigenvector which is associated to the largest volatility mode in the market i.e. largest eigenvalue of the realized correlation matrix. Interestingly, we notice that, whereas in the absence of feedback effects and using the parameters of Corollary 3.4, the first eigenvector should be proportional to the market vector $\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}$, due to the impact of the fund's rebalancing trades and in the case of a well-diversified fund (n large), the first eigenvector is also driven by the vector $\begin{pmatrix}\Delta\Phi_T^1\\ \vdots\\ \Delta\Phi_T^n\end{pmatrix}$ which represents the (time-weighted) rebalancing volumes by the large institutional investor. In scenarios where the fund trades significantly more in one asset than in others, its impact on returns will give more weight to this asset in the first eigenvalue of the correlation matrix.

4 Asset returns and fund performance

In this section, we study the impact of the fund's rebalancing strategy on asset and fund expected returns. Recall that, in the absence of feedback effects from the fund, the

	Theoretical eigenvalue Eq (14) vs Numerical eigenvalue	Fundamental eigenvalue vs Numerical eigenvalue
Average Error	10%	60%

Table 3: Average error for the largest eigenvector of the realized correlation matrix

benchmark return for asset i is m_i and hence the benchmark return for the institutional investor is $\sum_{i=1}^{n} x_i m_i$.

4.1 Asset expected returns

Following the same method as in Corollary 3.1, we can prove the following Proposition which gives the expansion at order one in $\|\Lambda\|$ of the expected return of each asset.

Proposition 4.1 *The (instantaneous) expected return of asset i at date t in the presence of feedback effects from the institutional investor investor is:*

$$\frac{b_i(P_t, V_t)}{P_t^i} = \underbrace{m_i}_{\substack{\text{benchmark return} \\ \text{for asset } i}} + \frac{\phi_t^i}{\phi_0^i}\Lambda_i(\underbrace{\sum_{l=1}^{n} x_l m_l}_{\substack{\text{benchmark return} \\ \text{for the fund}}} - m_i) + O(\|\Lambda\|^2) \tag{16}$$

where b_i, Λ and ϕ_t^i are defined respectively in Theorem 2.2, Eq (8) and Eq (4).

Trading by the institutional investor generates a systematic non fundamental component in the expected return of each asset. The nature of the fund's impact on the expected return of asset i depends on the difference between the benchmark expected return for asset i, m_i, and the benchmark expected return for the fund, $\sum_{l=1}^{n} x_l m_l$. When $m_i > \sum_{l=1}^{n} x_l m_l$, Eq (16) shows that $\frac{b_i(P_t,V_t)}{P_t^i} < m_i$ and *trading by the large institutional investor decreases (resp. increases) the expected return of assets whose benchmark expected returns are larger (resp. lower) than the benchmark return of the fund.* The fund's rebalancing strategy endogenously dampens the difference between expected returns of assets with large fundamental expected returns (compared to the benchmark) and those with low fundamental expected returns.

Proposition 4.1 enables us to compute the expected return for the large institutional investor and for other (small) investors.

Corollary 4.2 *At date t, the (instantaneous) expected returns for the large institutional investor and for a (small) investor holding a portion y_t^i of each asset i are given respectively by*

$$\sum_{i=1}^{n} x_i m_i + \sum_{i=1}^{n} x_i \Lambda_i \phi_t^i \left(\sum_{l=1}^{n} x_l m_l - m_i \right) + O(\|\Lambda\|^2)$$
$$\sum_{i=1}^{n} y_t^i m_i + \sum_{i=1}^{n} y_t^i \Lambda_i \phi_t^i \left(\sum_{l=1}^{n} x_l m_l - m_i \right) + O(\|\Lambda\|^2)$$

The expected return for the large institutional investor is lower (resp. higher) than its benchmark return $\sum_{i=1}^{n} x_i m_i$ if it overweighs (resp. underweighs) assets with large fundamental expected returns. *When the fund overweighs assets with large expected returns, it will generate, due to its own market impact, lower-than-expected returns.*

The expected return for a small fund with positions y_t is lower than the small fund's fundamental return $\sum_{i=1}^{n} y_t^i m_i$ when it overweighs assets with large fundamental expected returns (ie when y_t^i is large for assets i which verify $m_i > \sum_{l=1}^{n} x_l m_l$). *Whereas large institutional investors are considered as benchmarks by other investors, we show that, precisely due to the large market impact of such large institutional investors, other investors who try to beat the large fund by overweighting (resp. underweighting) assets with expected returns larger (resp. lower) than the large fund's benchmark will experience lower-than-expected returns.*

4.2 Optimal strategy and efficient frontier in a simple example

The following example shows how supposedly optimal strategies become sub-optimal due to the presence of large investors. When the size of the fund is significant, all investors (including the fund itself) have to take into account the impact of the large fund when choosing their allocations.

We consider the case of a market with $n = 2$ assets with zero fundamental correlation and identical fundamental volatility. We write:

$$\Sigma = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix} \quad \textit{and} \quad m = \begin{pmatrix} m_1 \\ m_2 \end{pmatrix}$$

and we assume, for example, that $m_1 > m_2$. The large institutional investor starts investing in this market and keeps a constant proportion of each asset in its portfolio, equal to 50% for each asset:

$$X = \begin{pmatrix} 50\% \\ 50\% \end{pmatrix}$$

Consider now a small fund investing in those two assets and choosing, at date t, an allocation y_t^1 in asset 1 and $y_t^2 = 1 - y_t^1$ of asset 2 by maximizing a mean-variance criteria. The small fund needs to estimate expected returns, variances and covariances and then calculates its allocation in each asset by solving a mean-variance criteria:

$$y_t^1 = \arg\max\{U_t(y); y \in \mathbb{R}\} \tag{17}$$

where

$$U_t(y) = y\mathbb{E}\left(return_{1,t}\right) + (1-y)\mathbb{E}\left(return_{2,t}\right)$$
$$-\gamma\left(y^2\mathbb{E}\left(variance_{1,t}\right) + (1-y)^2\,\mathbb{E}\left(variance_{2,t}\right) + 2y\left(1-y\right)\mathbb{E}\left(covariance_{1,2,t}\right)\right)$$

where γ is a parameter of risk aversion for the small fund.

In the presence of feedback effects from the rebalancing by the large institutional investor, using Proposition 3.1 and Proposition 4.1, we find that:

$$U_0(y) = ym_1 + (1-y)m_2 - \gamma\sigma^2\left(y^2 + (1-y)^2\right) + \tag{18}$$
$$+\left(\frac{y\Lambda_1 - (1-y)\Lambda_2}{2}(m_2 - m_1)\right) + \gamma\sigma^2\left(y^2\Lambda_1 + (1-y)^2\Lambda_2 - y(1-y)(\Lambda_1+\Lambda_2)\right) + O(\|\Lambda\|^2)$$

Thanks to Eq (18), the following Proposition follows immediately. Notice that we choose to focus on date 0 for clarity purpose only.

Proposition 4.3 *At date 0, the optimal allocation in asset* 1*, associated to the mean-variance criteria in Eq* (17)*, is given by:*

$$y_0^1 = \underbrace{\frac{1}{2} + \frac{m_1 - m_2}{4\gamma\sigma^2}}_{\substack{benchmark \\ optimal\ strategy}} + (\Lambda_1 + \Lambda_2)\frac{m_1 - m_2}{8\gamma\sigma^2} + \frac{\Lambda_1 - \Lambda_2}{4} + O(\|\Lambda\|^2)$$

When there are no feedback effects, the mean variance criteria gives an optimal proportion in each asset which is constant and hence implies that the small fund's strategy will be a fixed-mix strategy. We see that the larger the difference between the fundamental expected return of asset 1 and 2, the greater the allocation in asset 1 (recall that $m_1 > m_2$). When risk aversion goes to infinity ($\gamma \to \infty$), the optimal allocations do not depend on the assets' fundamental expected returns.

In the presence of feedback effects, the strategy $\frac{1}{2} + \frac{m_1 - m_2}{4\gamma\sigma^2}$ is no longer optimal. *The presence of large institutional investors generates a non-optimality for strategies which are optimal in the absence of feedback effects.* If the small fund does not take into account the market impact of the large fund, it will choose the benchmark optimal strategy $\frac{1}{2} + \frac{m_1 - m_2}{4\gamma\sigma^2}$ which is not optimal. On the contrary, if the small fund estimates the fundamentals of the market (from price series when the large investor was not trading in the market) and knows the strategy of the large institutional investor (which is realistic, for example, for large mutual funds whose strategies have to be disclosed), the small fund will be able to follow the strategy y_t^1 which is optimal for the mean-variance criteria in

Eq (17). Figure 3 shows that the efficient frontier is modified in the presence of feedback effects. We see that taking feedback effects into account enables the small investor to diminish the volatility of its portfolio for a given return. This stems from the fact that fixed-mix rebalancing induces a decrease of realized asset volatilities, as shown in Section 3.2, which can be used by an investor following the optimal strategy y_t^1 to build a less volatile portfolio. *Trading and risk-management decisions by investors need to take into account the allocations and sizes of large benchmark portfolios built by institutional investors.*

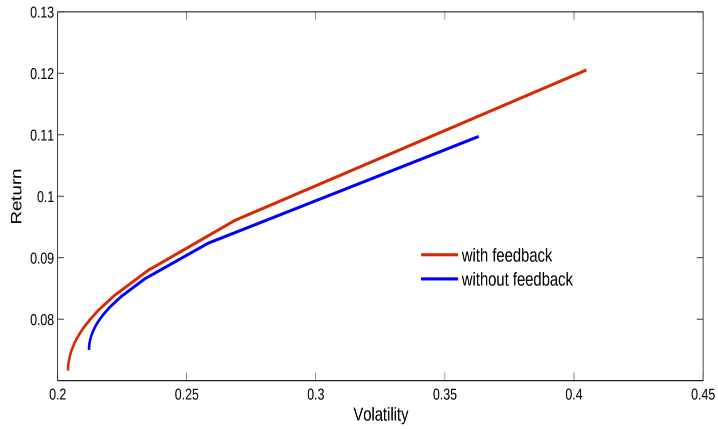

Figure 3: Efficient frontier with and without feedback effects

4.3 Trade-off between diversification and impact

The rebalancing impact described in the previous sections limits the benefits of investing in the fixed-mix strategy. Typically, large institutional investors prefer investing a significant portion of their wealth in such strategies, in order to enhance their returns, and keep a minimum portion in cash or money-market securities. We show that, *whereas it is indeed favorable for the fund to diversify by investing a portion of its wealth in a fixed-mix strategy, there exists a critical size for such an investment above which the impact of rebalancing generates losses for the fund.*

Consider that the fund holds a quantity C of cash and it can invest a portion $\alpha \in [0, 1]$ in the fixed-mix strategy and a portion $1 - \alpha$ at the risk free rate r (typically through T-Bills). Given Corollary 4.2, the expected return for the fund at date 0 is hence equal to:

$$(1-\alpha)r + \alpha \left(\sum_{l=1}^{n} x_l m_l + \sum_{i=1}^{n} x_i \Lambda_i (\sum_{l=1}^{n} x_l m_l - m_i) \right)$$

Typically, in order to generate a large return, the fund chooses a vector of allocations $(x_1, ..., x_n)$ for the fixed-mix strategy which overweights assets with large expected returns and hence, for all $\sum_{i=1}^{n} x_i \Lambda_i (\sum_{l=1}^{n} x_l m_l - m_i) < 0$. In addition, as the fund invests a portion α of its cash C in the fixed-mix strategy, we have $\Lambda_i = \alpha C \frac{x_i}{P_0^i D_i}$. As a consequence, the expected return for the fund can be written as

$$\alpha^2 \underbrace{\left(\sum_{i=1}^{n} x_i \frac{x_i C}{P_0^i D_i} (\sum_{l=1}^{n} x_l m_l - m_i) \right)}_{<0} + \alpha \underbrace{\left(\sum_{l=1}^{n} x_l m_l - r \right)}_{>0} + r \tag{19}$$

Eq (19) shows that the return of the expected return for the fund is not increasing in α.

There exists $\alpha^* = \frac{\sum_{l=1}^{n} x_l m_l - r}{2 \sum_{i=1}^{n} \frac{x_i^2 C}{P_0^i D_i} (m_i - \sum_{l=1}^{n} x_l m_l)}$ such that

- the return is increasing on $[0, \alpha^*]$: when the fund starts investing in the fixed-mix strategy, such investment actually increases its expected return. The more the fund invests in the fixed-mix strategy, as long as the portion allocated is lower than α^*, the larger the expected return for the fund.
- the return is decreasing on $[\alpha^*, 1]$: when the fund invests more than a portion α^* of its wealth in the fixed-mix strategy, its own rebalancing impact actually lowers its expected return and, the more it overweighs the fixed-mix strategy (above this weight of α^*), the lower the expected return for the fund.

The critical size for the investment in the fixed-mix strategy is equal to

$$\alpha^* C = \frac{\sum_{l=1}^{n} x_l m_l - r}{2 \sum_{i=1}^{n} \frac{x_i^2}{P_0^i D_i} (m_i - \sum_{l=1}^{n} x_l m_l)}$$

When the fund invests more than $\alpha^* C$ in the fixed-mix strategy, its performance will be diminished systematically by its own market impact.

5 Appendices

We denote

$$M_k = \begin{pmatrix} S_k \\ W_k \end{pmatrix} \in \left(\mathbb{R}_+^*\right)^{n+1} \tag{20}$$

and

$$Z_{k+1}^i = \Delta t (m_i - \frac{\Sigma_{i,i}}{2}) + \sqrt{\Delta} t \xi_{k+1}^i \tag{21}$$

As the large fund has long positions, we know that $x_i \geq 0$ for all $1 \leq i \leq n$. In addition, we have $\sum_{i=1}^{n} x_i = 1$.

5.1 Proof of Proposition 2.1

Let us show the following lemma which will imply Proposition 2.1.

Lemma 5.1 *There exists $\theta : (\mathbb{R}_+^*)^{n+1} \times \mathbb{R}^n \to (\mathbb{R}_+^*)^{n+1}$ such that for every compact set $K \subset (\mathbb{R}_+^*)^{n+1}$, there exists $\epsilon_K > 0$ such that θ is $\mathcal{C}^\infty$ on $K \times \mathcal{B}(0, \epsilon_K)$ and such that for all periods $k \geq 0$:*

$$\begin{pmatrix} S_{k+1} \\ W_{k+1} \end{pmatrix} = \begin{pmatrix} S_{k+1}^1 \\ \vdots \\ S_{k+1}^n \\ W_{k+1} \end{pmatrix} = \begin{pmatrix} \theta_1(S_k, W_k, Z_{k+1}) \\ \vdots \\ \theta_n(S_k, W_k, Z_{k+1}) \\ \theta_{n+1}(S_k, W_k, Z_{k+1}) \end{pmatrix}$$

and (S_k, W_k) verifies Equations 1, 2, and 3. Here Z_{k+1} is defined in Eq (21).

Proof Let $M_k \in (\mathbb{R}_+^*)^{n+1}$, defined in Eq. Eq (20), and ξ_{k+1} be given, thus fixing the value of $\phi_k^i = \frac{x_i M_k^{n+1}}{M_k^i}$. We can write Eq (3) as $M_{k+1}^i = A_i(M_k, Z_{k+1}) + \frac{1}{M_{k+1}^i} B_i(M_k, Z_{k+1}) M_{k+1}^{n+1}$ for $1 \leq i \leq n$, where A_i and B_i are defined on $(\mathbb{R}_+^*)^{n+1} \times \mathbb{R}^n$ by:

$$A_i(M, Z) = M^i \exp(Z^i) \left(1 - \frac{x_i M^{n+1}}{M^i D_i}\right) \tag{22}$$

$$B_i(M, Z) = M^i \exp(Z^i) \frac{x_i}{D_i} \tag{23}$$

which implies that $M_{k+1}^i = \frac{1}{2}\left(A_i(M_k, Z_{k+1}) + \sqrt{A_i^2(M_k, Z_{k+1}) + 4B_i(M_k, Z_{k+1}) M_{k+1}^{n+1}}\right)$. Reinjecting in Equation 2, we find that M_{k+1}^{n+1} is a fixed point of the function

$$v(x) = \frac{1}{2} \sum_{i=1}^n \frac{x_i M_k^{n+1}}{M_k^i} \left(A_i(M_k, Z_{k+1}) + \sqrt{A_i^2(M_k, Z_{k+1}) + 4B_i(M_k, Z_{k+1}) x}\right)$$

Let us show that v has a unique fixed point on $\mathbb{R}_+^*$.

Existence Given the expression of v, it is clear that for x large enough, $v(x) < x$. We then examine the three following possibilities:

- there exists i_0 such that $A_{i_0}(M_k, Z_{k+1}) > 0$, which implies that $v(0) > 0$ and, as v is a continuous function of x, that $v(x) > x$ for x small enough;
- there exists i_0 such that $A_{i_0}(M_k, Z_{k+1}) = 0$, which implies that

  $$v(x) \geq \frac{1}{2} \frac{x_i M_k^{n+1}}{M_k^{i_0}} \sqrt{4B_{i_0}(M_k, Z_{k+1}) x}$$

  which is strictly larger than x for x small enough;
- $\forall i \; A_i(M_k, Z_{k+1}) < 0$ which implies that $v(0) = 0$. Let us then calculate:

  $$v'(0) = \sum_{i=1}^n \frac{x_i M_k^{n+1}}{M_k^i} \frac{B_i(M_k, Z_{k+1})}{|A_i(M_k, Z_{k+1})|} = \sum_{i=1}^n \frac{x_i M_k^{n+1}}{M_k^i} \frac{\frac{x_i}{D_i}}{\frac{x_i M_k^{n+1}}{M_k^i D_i} - 1}$$

This implies that:

$$v'(0) > \sum_{i=1}^{n} \frac{x_i M_k^{n+1}}{M_k^i} \frac{\frac{x_i}{D_i}}{\frac{x_i M_k^{n+1}}{M_k^i D_i}} = \sum_{i=1}^{n} x_i = 1$$

assuring that $v(x) > x$ for x small enough.

As v is a continuous function of x, that $v(x) > x$ for large x and $v(x) < x$ for small x, there exists at least one fixed point $x_0 > 0$ such that $v(x_0) = x_0$.

Unicity Suppose that there exist two fixed points of function v, denoted a and b with $0 < a < b$. As v is concave, for $0 < x < a$ we have $\frac{v(a)-v(x)}{a-x} \geq \frac{v(b)-v(a)}{b-a} = 1$, meaning that $x \geq v(x)$ which is in contradiction with the fact that $x < v(x)$ for x sufficiently small. As a consequence, v cannot have more than one fixed point. The unique fixed point of v is $M_{k+1}^{n+1} > 0$ and we can deduce, for $1 \leq i \leq n$:

$$M_{k+1}^i = \frac{1}{2}\left(A_i(M_k, Z_{k+1}) + \sqrt{A_i^2(M_k, Z_{k+1}) + 4B_i(M_k, Z_{k+1})M_{k+1}^{n+1}}\right) > 0$$

This proves Proposition 2.1. We now denote $\psi : (\mathbb{R}_+^*)^{n+1} \times \mathbb{R}^n \times \mathbb{R}_+^* \to \mathbb{R}$ defined by:

$$\psi : (M, Z, x) \to x - \frac{1}{2}\sum_{i=1}^{n} \frac{x_i M^{n+1}}{M^i}\left(A_i(M, Z) + \sqrt{A_i^2(M, Z) + 4B_i(M, Z)x}\right)$$

where A_i and B_i are defined respectively in Eq (22) and Eq (23). ψ is $\mathcal{C}^\infty$. Furthermore, we have $\psi(M, 0, M^{n+1}) = 0$ and $A_i(M, 0) = M^i\left(1 - \frac{x_i M^{n+1}}{M^i D_i}\right)$ and $B_i(M, 0) = M^i \frac{x_i}{D_i}$ according to Eq (22) and Eq (23), which implies that

$$A_i^2(M, 0) + 4B_i(M, 0)M^{n+1} = \left(M^i\left(1 + \frac{x_i M^{n+1}}{M^i D_i}\right)\right)^2 \tag{24}$$

Hence we find that:

$$\begin{aligned}
\frac{\partial\psi}{\partial x}(M, 0, M^{n+1}) &= 1 - \sum_{i=1}^{n} \frac{x_i M^{n+1}}{M^i} \frac{B_i(M, 0)}{\sqrt{A_i^2(M, 0) + 4B_i(M, 0)M^{n+1}}} \\
&= 1 - \sum_{i=1}^{n} x_i \times \frac{\frac{x_i M^{n+1}}{M^i}}{D_i + \frac{x_i M^{n+1}}{M^i}} \\
&= \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i \times \frac{\frac{x_i M^{n+1}}{M^i}}{D_i + \frac{x_i M^{n+1}}{M^i}} \\
\frac{\partial\psi}{\partial x}(M, 0, M^{n+1}) &= \sum_{i=1}^{n} \frac{x_i}{1 + \frac{x_i M^{n+1}}{D_i M^i}} > 0 \qquad (25)
\end{aligned}$$

As a consequence, if K is a compact set of $\left(\mathbb{R}_+^*\right)^{n+1}$, the implicit function theorem states that there exists $\epsilon_K > 0$ and θ_{n+1} which is $\mathcal{C}^\infty$ on $K \times \mathcal{B}(0, \epsilon_K)$ such that:

$$\psi(M, Z, \theta_{n+1}(M, Z)) = 0 \tag{26}$$

and then, we deduce, for $1 \leq i \leq n$,

$$\theta_i(M,Z) = \frac{1}{2}\left(A_i(M,Z) + \sqrt{A_i^2(M,Z) + 4B_i(M,Z)\theta_{n+1}(M,Z))}\right)$$

and θ_i is $\mathcal{C}^\infty$ on $K \times \mathcal{B}(0, \epsilon_K)$. This concludes the proof for the existence and smoothness of θ.

5.2 Proof of Theorem 2.2

Assumption 5.2 *There exists $\eta > 0$ such that:*

$$\mathbb{E}(\|\exp(\eta\xi)\|) < \infty \quad \textit{and} \quad \mathbb{E}(\|\xi\|^{\eta+4}) < \infty$$

Define a (resp., b) a $\mathcal{M}_{n+1\times n}(\mathbb{R})$-valued (resp. $\mathbb{R}^{n+1}$-valued) mapping such that

$$a_{i,j}(S,W) = \sum_{l=1}^{n} \frac{\partial \theta_i}{\partial z_l}(S,W,0) \times A_{l,j} \tag{27}$$

$$b_i(S,W) = \sum_{j=1}^{n} \frac{\partial \theta_i}{\partial z_j}(S,W,0)\overline{m_j} + \frac{1}{2}\sum_{j,l=1}^{n} \frac{\partial^2 \theta_i}{\partial z_j \partial z_l}(S,W,0)\Sigma_{j,l} \tag{28}$$

where θ is defined in Lemma 5.1, $\overline{m}_i = m_i - \frac{\Sigma_{i,i}}{2}$ and A is a square-root of the fundamental covariance matrix: $A\ {}^tA = \Sigma$.

Using Lemma 5.1 and (Cont and Wagalath, 2013b, Section 6.1), we directly find the following Lemma:

Lemma 5.3 *Under Assumption 5.2, for all $\epsilon > 0$, $r > 0$:*

$$\lim_{\Delta t \to 0} \sup_{\|M\| \leq r} \frac{1}{\Delta t}\mathbb{P}(\|M_{k+1} - M_k\| \geq \epsilon | M_k = M) = 0 \tag{29}$$

$$\lim_{\Delta t \to 0} \sup_{\|M\| \leq r} \left\| \frac{1}{\Delta t}\mathbb{E}(M_{k+1} - M_k | M_k = M) - b(M) \right\| = 0 \tag{30}$$

$$\lim_{\Delta t \to 0} \sup_{\|M\| \leq r} \left\| \frac{1}{\Delta t}\mathbb{E}[(M_{k+1} - M_k)(M_{k+1} - M_k)^t | M_k = M] - aa^t(M) \right\| = 0 \tag{31}$$

where a and b are defined respectively in Eq (27) and Eq (28).

The following lemmas are a direct consequence of the implicit function theorem and enable to compute a and b explicitely.

Lemma 5.4 *For $1 \leq l, i \leq n$:*

$$\frac{\partial \theta_i}{\partial z_l}(M,0) = \frac{M^i}{1 + \frac{x_i M^{n+1}}{M^i D_i}} \left(\delta_{i,l} + \frac{\frac{x_i M^{n+1}}{M^i D_i}}{\sum_{j=1}^{n} \frac{x_j}{1 + \frac{x_j M^{n+1}}{M^j D_j}}} \times \frac{x_l}{1 + \frac{x_l M^{n+1}}{M^l D_l}} \right) \tag{32}$$

$$\frac{\partial \theta_{n+1}}{\partial z_l}(M,0) = \frac{M^{n+1}}{\sum_{j=1}^{n} \frac{x_j}{1 + \frac{x_j M^{n+1}}{M^j D_j}}} \frac{x_l}{1 + \frac{x_l M^{n+1}}{M^l D_l}} \tag{33}$$

Proof θ_{n+1} is defined implicitely by $\psi(M,Z,\theta_{n+1}(M,Z))=0$, where ψ is given in Eq (**??**). The implicit function theorem gives, for $1\le l\le n$,

$$\frac{\partial\theta_{n+1}}{\partial z_l}(M,Z)=\frac{-\frac{\partial\psi}{\partial z_l}(M,Z,\theta_{n+1}(M,Z))}{\frac{\partial\psi}{\partial x}(M,Z,\theta_{n+1}(M,Z))} \tag{34}$$

Given the expression for ψ given in Eq (**??**), we find that

$$\frac{\partial\psi}{\partial z_l}(M,Z,\theta_{n+1}(M,Z))=\frac{-1}{2}\frac{x_lM^{n+1}}{M^l}\left(A_l(M,Z)+\frac{A_l^2(M,Z)+2B_l(M,Z)x}{\sqrt{A_l^2(M,Z)+4B_l(M,Z)x}}\right) \tag{35}$$

and

$$\frac{\partial\psi}{\partial x}(M,Z,\theta_{n+1}(M,Z))=1-\sum_{i=1}^{n}\frac{x_iM^{n+1}}{M^i}\frac{B_i(M,Z)}{\sqrt{A_i^2(M,Z)+4B_i(M,Z)x}} \tag{36}$$

Using Eq (22) and Eq (23), we find that

$$A_l^2(M,Z)+2B_l(M,Z)M^{n+1}=(M^l)^2\left(1+\left(\frac{x_lM^{n+1}}{M^lD_l}\right)^2\right) \tag{37}$$

Given Eq (24) and the fact that $\theta_{n+1}(M,0)=M^{n+1}$, we find that

$$\frac{\partial\psi}{\partial z_l}(M,0,\theta_{n+1}(M,0))=\frac{-x_lM^{n+1}}{1+\frac{x_lM^{n+1}}{M^lD_l}} \tag{38}$$

Using Eq (34), Eq (38) and Eq (25), we find that

$$\frac{\partial\theta_{n+1}}{\partial z_l}(M,0)=\frac{M^{n+1}}{\displaystyle\sum_{j=1}^{n}\frac{x_j}{1+\frac{x_jM^{n+1}}{M^jD_j}}}\frac{x_l}{1+\frac{x_lM^{n+1}}{M^lD_l}}$$

Let us now calculate, for $1\le i\le n$ and $1\le l\le n$, $\frac{\partial\theta_i}{\partial z_l}(M,0)$. Given the definition of θ_i, we remark that:

$$\frac{\partial\theta_i}{\partial z_l}(M,Z)=-\frac{\partial\psi}{\partial z_l}(M,Z,\theta_{n+1}(M,Z))\frac{\delta_{i,l}}{\frac{x_iM^{n+1}}{M^i}}+\frac{B_i(M,Z)\frac{\partial\theta_{n+1}}{\partial z_l}(M,Z)}{\sqrt{A_l^2(M,Z)+4B_l(M,Z)x}} \tag{39}$$

which, evaluated in 0 and given the expression for $\frac{\partial\theta_{n+1}}{\partial z_l}$ and Eq (24) gives:

$$\frac{\partial\theta_i}{\partial z_l}(M,0)=\frac{M^i}{1+\frac{x_iM^{n+1}}{M^iD_i}}\left(\delta_{i,l}+\frac{\frac{x_iM^{n+1}}{M^iD_i}}{\displaystyle\sum_{j=1}^{n}\frac{x_j}{1+\frac{x_jM^{n+1}}{M^jD_j}}}\times\frac{x_l}{1+\frac{x_lM^{n+1}}{M^lD_l}}\right)$$

Lemma 5.5 *For* $1\le j,l\le n$*:*

$$\frac{\partial^2\theta_{n+1}}{\partial z_j\partial z_l}(M,0)\frac{1}{M^{n+1}}\sum_{1\le p\le n}\frac{x_p}{1+\frac{x_pM^{n+1}}{D_pM^p}}=\frac{\delta_{j,l}x_l}{1+\frac{x_lM^{n+1}}{D_lM^l}}\left(1-\frac{2\left(\frac{x_lM^{n+1}}{D_lM^l}\right)^2}{\left(1+\frac{x_lM^{n+1}}{D_lM^l}\right)^2}\right) \tag{40}$$

$$+\frac{2x_jx_l}{\left(1+\frac{x_lM^{n+1}}{D_lM^l}\right)\left(1+\frac{x_jM^{n+1}}{D_jM^j}\right)}\frac{1}{\sum_{1\le p\le n}\frac{x_p}{1+\frac{x_pM^{n+1}}{D_pM^p}}}\left(\frac{\left(\frac{x_lM^{n+1}}{D_lM^l}\right)^2}{\left(1+\frac{x_lM^{n+1}}{D_lM^l}\right)^2}+\frac{\left(\frac{x_jM^{n+1}}{D_jM^j}\right)^2}{\left(1+\frac{x_jM^{n+1}}{D_jM^j}\right)^2}\right)$$

$$-\frac{2x_jx_l}{\left(1+\frac{x_lM^{n+1}}{D_lM^l}\right)\left(1+\frac{x_jM^{n+1}}{D_jM^j}\right)}\frac{1}{\left(\sum_{1\le p\le n}\frac{x_p}{1+\frac{x_pM^{n+1}}{D_pM^p}}\right)^2}\sum_{1\le p\le n}\frac{x_p\left(\frac{x_pM^{n+1}}{D_pM^p}\right)^2}{\left(1+\frac{x_pM^{n+1}}{D_pM^p}\right)^3}$$

$$\frac{\partial^2\theta_i}{\partial z_j\partial z_l}(M,0,M^{n+1})=\delta_{i,j}\delta_{j,l}\frac{M^i}{1+\frac{x_iM^{n+1}}{D_iM^i}}\left(1-\frac{2\left(\frac{x_lM^{n+1}}{D_lM^l}\right)^2}{\left(1+\frac{x_lM^{n+1}}{D_lM^l}\right)^2}\right) \tag{41}$$

$$+2M^i\frac{\left(\frac{x_iM^{n+1}}{D_iM^i}\right)^2}{\left(1+\frac{x_iM^{n+1}}{D_iM^i}\right)^3}\frac{1}{\sum_{1\le p\le n}\frac{x_p}{1+\frac{x_pM^{n+1}}{D_pM^p}}}\left(\delta_{i,l}\frac{x_j}{\left(1+\frac{x_jM^{n+1}}{D_jM^j}\right)}+\delta_{i,j}\frac{x_l}{\left(1+\frac{x_lM^{n+1}}{D_lM^l}\right)}\right)$$

$$-2M_i\frac{2\left(\frac{x_lM^{n+1}}{D_lM^l}\right)^2}{\left(1+\frac{x_lM^{n+1}}{D_lM^l}\right)^3}\frac{\frac{x_j}{1+\frac{x_jM^{n+1}}{D_jM^j}}\frac{x_l}{1+\frac{x_lM^{n+1}}{D_lM^l}}}{\left(\sum_{1\le p\le n}\frac{x_p}{1+\frac{x_pM^{n+1}}{D_pM^p}}\right)^2}+\frac{\frac{x_i}{D_i}}{1+\frac{x_iM^{n+1}}{D_iM^i}}\frac{\partial^2\theta_{n+1}}{\partial z_j\partial z_l}(M,0)$$

Proof We first calculate for $1\le j,l\le n$ $\frac{\partial^2\theta_{n+1}}{\partial z_j\partial z_l}(M,0)$. Deriving Eq (35) with respect to z_l and z_j gives the following equation:

$$\frac{\partial^2\psi}{\partial z_j\partial z_l}(M,Z,\theta_{n+1}(M,Z))+\frac{\partial^2\psi}{\partial x\partial z_l}(M,Z,\theta_{n+1}(M,Z))\frac{\partial\theta_{n+1}}{\partial z_j}(M,Z) \tag{42}$$

$$+\frac{\partial^2\psi}{\partial x\partial z_j}(M,Z,\theta_{n+1}(M,Z))\frac{\partial\theta_{n+1}}{\partial z_l}(M,Z)+\frac{\partial^2\psi}{\partial x^2}(M,Z,\theta_{n+1}(M,Z))\frac{\theta_{n+1}}{\partial z_j}(M,Z)\frac{\theta_{n+1}}{\partial z_l}(M,Z)$$

$$+\frac{\partial\psi}{\partial x}(M,Z,\theta_{n+1}(M,Z))\frac{\partial^2\theta_{n+1}}{\partial z_j\partial z_l}(M,Z)=0$$

Considering Eq (35), we find that if $l\neq j$, then $\frac{\partial^2\psi}{\partial z_j\partial z_l}(M,Z,\theta_{n+1}(M,Z))=0$. Deriving Eq (35) with respect to z_l, we find that

$$\frac{\partial^2\psi}{\partial z_l^2}(M,Z,x)=-\frac{1}{2}\frac{x_lM^{n+1}}{M^l}\left(A_l(M,Z)+\frac{2A_l^2(M,Z)+2B_l(M,Z)x}{\sqrt{A_l^2(M,Z)+4B_l(M,Z)x}}-\frac{(A_l^2(M,Z)+2B_l(M,Z)x)^2}{(A_l^2(M,Z)+4B_l(M,Z)x)^{\frac{3}{2}}}\right)$$

$$=\frac{\partial\psi}{\partial z_l}(M,Z,x)+\frac{2x_lM^{n+1}}{M^l}\frac{B_l^2(M,Z)x^2}{(A_l^2(M,Z)+4B_l(M,Z)x)^{\frac{3}{2}}}$$

Evaluating this equation in 0 and using Eq (38) and Eq (24), we find that

$$\frac{\partial^2\psi}{\partial z_j\partial z_l}(M,0,M^{n+1})=\delta_{j,l}\frac{x_lM^{n+1}}{\left(1+\frac{x_lM^{n+1}}{D_lM^l}\right)^3}\left(\left(1-\frac{x_lM^{n+1}}{D_lM^l}\right)^2-1\right) \tag{43}$$

Deriving Eq (35) with respect to x, we find that

$$\frac{\partial^2\psi}{\partial z_l \partial x}(M,Z,x) = -\frac{2x_l M^{n+1}(M,Z)}{M^l}\frac{B_l^2(M,Z)x}{(A_l^2(M,Z)+4B_l(M,Z)x)^{\frac{3}{2}}}$$

which, evaluated in 0, gives

$$\frac{\partial^2\psi}{\partial z_l \partial x}(M,0,M^{n+1}) = -2x_l\left(\frac{M^{n+1}x_l}{M^l D_l}\right)^2\frac{1}{\left(1+\frac{M^{n+1}x_l}{M^l D_l}\right)^3} \quad (44)$$

Differentiating Eq (36) with respect to x, we find that

$$\frac{\partial^2\psi}{\partial x^2}(M,Z,x) = 2\sum_{1\le p\le n}\frac{x_p M^{n+1}}{M^p}\frac{B_p^2(M,Z)}{(A_p^2(M,Z)+4B_p(M,Z)x)^{\frac{3}{2}}}$$

which implies that

$$\frac{\partial^2\psi}{\partial x^2}(M,0,M^{n+1}) = \frac{2}{M^{n+1}}\sum_{1\le p\le n}\frac{x_p\left(\frac{x_p M^{n+1}}{D_p M^p}\right)^2}{\left(1+\frac{x_p M^{n+1}}{D_p M^p}\right)^3} \quad (45)$$

Using Eq (32), Eq (33), Eq (43), Eq (44) and Eq (45), the relationship given in Eq (42) gives Eq (40) of Lemma 5.5.

Using Eq (39), we can calculate the second order derivative of θ_i for $1 \le i \le n$. For $1 \le j,l \le n$:

$$\begin{aligned}\frac{\partial^2\theta_i}{\partial z_j \partial z_l}(M,Z,\theta_{n+1}(M,Z)) &= -\frac{\delta_{i,l}}{\frac{x_i M^{n+1}}{M^i}}\left(\frac{\partial^2\psi}{\partial z_j \partial z_l}(M,Z,\theta_{n+1}(M,Z))\right)\\ &-\frac{\delta_{i,l}}{\frac{x_i M^{n+1}}{M^i}}\left(\frac{\partial^2\psi}{\partial x \partial z_l}(M,Z,\theta_{n+1}(M,Z))\frac{\partial\theta_{n+1}}{\partial z_j}(M,Z)\right)\\ &+\delta_{i,j}\frac{\partial\theta_{n+1}}{\partial z_l}(M,Z)\frac{2B_i^2(M,Z)\theta_{n+1}(M,Z)}{(A_i^2(M,Z)+4B_i(M,Z)\theta_{n+1}(M,Z))^{\frac{3}{2}}}\end{aligned}$$

$$-2\frac{B_i^2(M,Z)\frac{\partial\theta_{n+1}}{\partial z_j}(M,Z)\frac{\partial\theta_{n+1}}{\partial z_l}(M,Z)}{(A_i^2(M,Z)+4B_i(M,Z)\theta_{n+1}(M,Z))^{\frac{3}{2}}}+\frac{B_i(M,Z)}{\sqrt{A_i^2(M,Z)+4B_i(M,Z)\theta_{n+1}(M,Z)}}\frac{\partial^2\theta_{n+1}}{\partial z_j \partial z_l}(M,Z)$$

which, for $Z=0$, gives Eq (41).

Define the differential operator $G : C_0^\infty(\mathbb{R}^{n+1}) \mapsto C_0^\infty(\mathbb{R}^{n+1})$ by

$$Gh(x) = \frac{1}{2}\sum_{1\le i,j\le n}(aa^t)_{i,j}(x)\partial_i\partial_j h + \sum_{i=1}^{n} b_i(x)\partial_i h$$

where a and b are defined in Eq (27) and Eq (28) respectively. a and b are continous and Lemmas 5.4 and 5.5 show that for all $x \in \mathbb{R}^{n+1}$, $\|a(x)\| + \|b(x)\| \le K\|x\|$. (Ethier and Kurtz, 1986, Theorem 2.6, Ch.8) states that the martingale problem for (G, δ_{S_0,W_0}) is well-posed. So, by Lemma 5.3 and (Ethier and Kurtz, 1986, Theorem 4.2, Ch.7), this implies that the process $(S_{\lfloor \frac{t}{\Delta t}\rfloor}, W_{\lfloor \frac{t}{\Delta t}\rfloor})$ converges in distribution to the solution $(\mathbb{P}, (P_t, V_t)_{t\ge 0})$ of the martingale problem for (G, δ_{S_0,W_0}) when $\Delta t \to 0$.

Furthermore, as a and b are $\mathcal{C}^\infty$, they are locally Lipschitz and hence, by (Ikeda and Watanabe, 1981, Theorem 3.1, Ch.4), the solution of the martingale problem for (G, δ_{S_0,W_0}) is the unique strong solution of the stochastic differential equation given in Theorem 2.2.

5.3 Proof of Proposition 3.1

We know that

$$C^{i,j}_{[0,T]} = \frac{[\ln P^i, \ln P^j]_T}{T} = \frac{1}{T}\int_0^T c^{i,j}_s \, \mathrm{d}s$$

where c_s is the derivative of the quadratic covariation and corresponds to the instantaneous covariance matrix of asset returns. By direct computation from Theorem 2.2 and Lemma 5.4, we find that

$$\begin{aligned} c_t = {} & \Sigma + \Sigma({}^t\Gamma_t - I_n)F_t + F_t(\Gamma_t - I_n)\Sigma \\ & + F_t(\Gamma_t - I_n)\Sigma({}^t\Gamma_t - I_n)F_t \end{aligned} \tag{46}$$

where F_t and Γ_t are $n \times n$ matrices such that F_t is diagonal with i-th term equal to

$$F^{i,i}_t = \frac{\Phi^i_t \Lambda_i}{1 + \Phi^i_t \Lambda_i} \quad and \quad \Gamma^{i,j}_t = \left(\sum_{1 \le p \le n} \frac{x_p}{1 + \Phi^p_t \Lambda_p} \right)^{-1} \frac{x_j}{1 + \Phi^j_t \Lambda_j} \tag{47}$$

with $\Phi^i_t = \frac{\phi^i_t}{\phi^i_0}$.This implies that:

$$\begin{aligned} C_{[0,T]} = {} & \Sigma + \frac{1}{T}\int_0^T \big(\Sigma({}^t\Gamma_s - I_n)F_s + F_s(\Gamma_s - I_n)\Sigma\big)\, ds \\ & + \frac{1}{T}\int_0^T \big(F_s(\Gamma_s - I_n)\Sigma({}^t\Gamma_s - I_n)F_s\big) ds \end{aligned} \tag{48}$$

where F_s and Γ_s are defined in Eq (47). We study the expression of $C_{[0,T]}$ given in Eq (48). Let us start with the term

$$\frac{1}{T}\int_0^T (F_s(\Gamma_s - I_n)\Sigma)\, ds$$

We have

$$[F_s(\Gamma_s - I_n)]_{i,k} = \frac{\Lambda_i \Phi^i_s}{1 + \Phi^i_s \Lambda_i} \left(\left(\sum_{1 \le p \le n} \frac{x_p}{1 + \Phi^p_s \Lambda_p} \right)^{-1} \frac{x_k}{1 + \Phi_k(s)\Lambda_k} - \delta_{i,k} \right)$$

which implies that

$$\int_0^T [F_s(\Gamma_s - I_n)\Sigma]_{i,j} ds = \int_0^T \sum_{1 \le k \le n} [F_s(\Gamma_s - I_n)]_{i,k} \Sigma_{k,j} ds$$

$$= -\int_0^T \frac{\Lambda_i \Phi^i_s}{1 + \Phi^i_s \Lambda_i} \Sigma_{i,j} ds + \int_0^T \frac{\Lambda_i \Phi^i_s}{1 + \Phi^i_s \Lambda_i} \sum_{1 \le k \le n} \left(\sum_{1 \le p \le n} \frac{x_p}{1 + \Phi^p_s \Lambda_p} \right)^{-1} \frac{x_k}{1 + \Phi_k(s)\Lambda_k} \Sigma_{k,j} ds$$

We then remark that

$$\left| \int_0^T \frac{\Lambda_i \Phi^i_s}{1 + \Phi^i_s \Lambda_i} \Sigma_{i,j} ds - \int_0^T \Lambda_i \Phi^i_s \Sigma_{i,j} ds \right| \le |\Sigma_{i,j}| \, \Lambda_i^2 \int_0^T \frac{(\Phi^i_s)^2}{1 + \Phi^i_s \Lambda_i} ds$$

and hence

$$\left|\int_0^T \frac{\Lambda_i\Phi_s^i}{1+\Phi_s^i\Lambda_i}\Sigma_{i,j}ds - \int_0^T \Lambda_i\Phi_s^i\Sigma_{i,j}ds\right| \le |\Sigma_{i,j}|\,\Lambda_i^2\int_0^T (\Phi_s^i)^2 ds \tag{49}$$

Furthermore, we remark that:

$$\left|\int_0^T \frac{\Lambda_i\Phi_s^i}{1+\Phi_s^i\Lambda_i}\sum_{1\le k\le n}\left(\sum_{1\le p\le n}\frac{x_p}{1+\Phi_s^p\Lambda_p}\right)^{-1}\frac{x_k}{1+\Phi_k(s)\Lambda_k}\Sigma_{k,j}ds - \int_0^T \Lambda_i\Phi_s^i\sum_{1\le k\le n}x_k\Sigma_{k,j}ds\right|$$

$$\le \sum_{k=1}^n x_k\,|\Sigma_{k,j}|\int_0^T \Phi_s^i\Lambda_i\left|\frac{1}{1+\Phi_s^i\Lambda_i}\frac{1}{1+\Phi_k(s)\Lambda_k}\left(\sum_{1\le p\le n}\frac{x_p}{1+\Phi_s^p\Lambda_p}\right)^{-1} - 1\right|ds$$

$$\le \sum_{k=1}^n x_k\,|\Sigma_{k,j}|\int_0^T \frac{\Phi_s^i\Lambda_i}{(1+\Phi_k(s)\Lambda_k)\,(1+\Phi_s^i\Lambda_i)}\left|\left(\sum_{1\le p\le n}\frac{x_p}{1+\Phi_s^p\Lambda_p}\right)^{-1} - (1+\Phi_k(s)\Lambda_k)\left(1+\Phi_s^i\Lambda_i\right)\right|ds$$

$$\le \sum_{k=1}^n x_k\,|\Sigma_{k,j}|\int_0^T \Phi_s^i\Lambda_i\left|\left(\sum_{1\le p\le n}\frac{x_p}{1+\Phi_s^p\Lambda_p}\right)^{-1} - (1+\Phi_k(s)\Lambda_k)\left(1+\Phi_s^i\Lambda_i\right)\right|ds$$

$$\le \sum_{k=1}^n x_k\,|\Sigma_{k,j}|\int_0^T \Phi_s^i\Lambda_i\left|\Phi_s^i\Lambda_i+\Phi_k(s)\Lambda_k+\Phi_s^i\Phi_k(s)\Lambda_i\Lambda_k+\left(\frac{\sum_{1\le p\le n}\frac{x_p}{1+\Phi_s^p\Lambda_p}-1}{\sum_{1\le p\le n}\frac{x_p}{1+\Phi_s^p\Lambda_p}}\right)\right|ds$$

$$\le \sum_{k=1}^n x_k\,|\Sigma_{k,j}|\int_0^T \Phi_s^i\Lambda_i\left|\Phi_s^i\Lambda_i+\Phi_k(s)\Lambda_k+\Phi_s^i\Phi_k(s)\Lambda_i\Lambda_k+\left(\frac{\sum_{1\le p\le n}\frac{x_p}{1+\Phi_s^p\Lambda_p}-\sum_{1\le p\le n}x_p}{\sum_{1\le p\le n}\frac{x_p}{1+\Phi_s^p\Lambda_p}}\right)\right|ds$$

$$\le \sum_{k=1}^n x_k\,|\Sigma_{k,j}|\int_0^T \Phi_s^i\Lambda_i\left(\Phi_s^i\Lambda_i+\Phi_k(s)\Lambda_k+\Phi_s^i\Phi_s^k\Lambda_i\Lambda_k-\left(\frac{\sum_{1\le p\le n}\frac{x_p\Phi_s^p\Lambda_p}{1+\Phi_s^p\Lambda_p}}{\sum_{1\le p\le n}\frac{x_p}{1+\Phi_s^p\Lambda_p}}\right)\right)ds$$

$$\le \sum_{k=1}^n x_k\,|\Sigma_{k,j}|\int_0^T \Phi_s^i\Lambda_i\left|\Phi_s^i\Lambda_i+\Phi_s^k\Lambda_k+\Phi_s^i\Phi_s^k\Lambda_i\Lambda_k+\sum_{1\le p\le n}x_p\Phi_s^p\Lambda_p\sum_{1\le p\le n}x_p(1+\Phi_s^p\Lambda_p)\right|ds \tag{50}$$

where we used that $\sum_{p=1}^n x_p = 1$ and that for strictly positive real numbers $(y_i)_{1\le i\le n}$, we have the convexity inequality $\left(\sum_{p=1}^n \frac{x_p}{y_p}\right)^{-1} \le \sum_{p=1}^n x_p y_p$.

Given Lemmas 5.4 and 5.5, we find that, for $1 \le i,k \le n$, $\frac{a_{i,k}(P_t,V_t)}{P_t^i}$, $\frac{a_{n+1,k}(P_t,V_t)}{V_t}$, $\frac{b_i(P_t,V_t)}{P_t^i}$ and $\frac{b_{n+1}(P_t,V_t)}{V_t}$, defined in Eq (27) and Eq (28), are bounded uniformly in Λ. As a consequence, by applying Itô's formula to $(\phi(t)^i)^p(\phi(t)^j)^q(\phi(t)^k)^r$ for $p,q,r \ge 0$, we find that $\mathbb{E}((\phi(t)^i)^p(\phi(t)^j)^q(\phi(t)^k)^r) \le K\exp(Ct)$ where C does not depend on Λ.

Given that $\Phi_s^i = \frac{\Phi_s^i}{\phi_0^i}$ and using Eq (49), we find that

$$\int_0^T \frac{\Lambda_i \Phi_s^i}{1+\Phi_s^i \Lambda_i} \Sigma_{i,j} ds = \int_0^T \Lambda_i \Phi_s^i \Sigma_{i,j} ds + O(\|\Lambda\|^2)$$

where $\mathbb{E}\left(\frac{O(\|\Lambda\|^2)}{\|\Lambda\|^2}\right)$ is bounded when Λ goes to zero.

Similarly, using Eq (50), we find that

$$\int_0^T \frac{\Lambda_i \Phi_s^i}{1+\Phi_s^i \Lambda_i} \sum_{k=1}^n \left(\sum_{p=1}^n \frac{x_p}{1+\Phi_s^p \Lambda_p} \right)^{-1} \frac{x_k}{1+\Phi_s^k \Lambda_k} \Sigma_{k,j} ds = \int_0^T \Lambda_i \Phi_s^i \sum_{k=1}^n x_k \Sigma_{k,j} ds + O(\|\Lambda\|^2)$$

We then use the same methodology to study the other terms of $C_{[0,T]}$ given in Eq (48). We conclude this proof by using Ito's formula and the fact that $\Phi_0^i = \frac{\phi_0^i}{\phi_0^i} = 1$, which gives the relationship

$$1 + \int_0^T \left(1 - \frac{s}{T}\right) d\Phi_s^i = \frac{1}{T} \int_0^T \Phi_s^i ds = \Delta \Phi_T^i$$

which leads to the decomposition of $C_{[0,T]}$ given in Proposition 3.1. The decomposition of $R_[0,T]$ follows directly, as $R_[0,T]^{i,j} = \frac{C^{i,j}_{[0,T]}}{\left(C_{i,i}([0,T]) C^{j,j}_{[0,T]}\right)^{\frac{1}{2}}}$.

5.4 Proof of Corollary 3.4

Proof In our example, the fundamental correlation matrix $\tilde{R}(0)$ is such that $[\tilde{R}(0)]_{i,i} = 1$ and $[\tilde{R}(0)]_{i,j} = \rho$ for $i \neq j$. As a consequence, it has a simple eigenvalue $v_1(0) = 1 + (n-1)\rho$, associated to the eigenvector $\psi_1(0) = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$, and an eigenvalue of order $n-1$: $v_k(0) = 1 - \rho$ for $2 \leq k \leq n$, associated to eigenvectors $(\psi_k(0))_{2 \leq k \leq n}$, which form an orthonormal basis of the hyperplan of $\mathbb{R}^n$: $\mathcal{H} = \{z \in \mathbb{R}^n s.t. \sum_{i=1}^n z_i = 0\}$. As the largest eigenvalue of the fundamental correlation matrix, $v_1(0)$, is a simple eigenvalue, we can use Proposition 3.3 to compute the largest eigenvalue and the first eigenvector of the realized correlation matrix of returns in the presence of feedback effects.

Given the parameters of our example, we find that for $i \neq j$

$$[\nabla R]_{i,j} = \overline{\Lambda}(1-\rho)(\rho + \frac{1-\rho}{n}) \left(2 + \int_0^T \left(1 - \frac{s}{T}\right) d\tilde{\Phi}_i(s) + \int_0^T \left(1 - \frac{s}{T}\right) d\tilde{\Phi}_j(s) \right)$$

We hence find the expression for $v(\overline{\Lambda})$ given in Corollary 3.4 by direct computation of Proposition 3.3.

Proposition 3.3 states that the first eigenvalue of the realized correlation matrix of returns in the presence of feedback effects is given by:

$$\psi_1 = \overline{\psi}_1 + \frac{1}{n\rho} \sum_{k=2}^n \left({}^t\overline{\psi}_k \nabla R \overline{\psi}_1 \right) \overline{\psi}_k + o(\overline{\Lambda})$$

as $v_1(0) - v_k(0) = n\rho$ for $k \geq 2$. Given the fact that $(\overline{\psi}_k)_{2\leq k\leq n}$ is an orthonormal basis of $\mathcal{H} = \{z \in \mathbb{R}^n s.t. \sum_{i=1}^{n} z_i = 0\}$, the term $\sum_{k=2}^{n} \left({}^t\overline{\psi}_k \nabla R\overline{\psi}_1\right) \overline{\psi}_k$ is the orthogonal projection of $\nabla R\overline{\psi}_1$ on $\mathcal{H}$. Given the equation defining $\mathcal{H}$, this orthogonal projection $p_{\mathcal{H}}$ on $\mathcal{H}$ is expressed as

$$p_{\mathcal{H}}(z) = \begin{pmatrix} z_1 - \frac{1}{n} \sum\limits_{1\leq i\leq n} z_i \\ \vdots \\ z_n - \frac{1}{n} \sum\limits_{1\leq i\leq n} z_i \end{pmatrix}$$

As $\left[\nabla R\overline{\psi}_1\right]_i = \frac{1}{\sqrt{n}}\overline{\Lambda}(1-\rho)(\rho + \frac{1-\rho}{n}) \sum_{j\neq i} \left(2 + \Delta\Phi_T^i + \Delta\Phi_T^j\right)$, and given the expression for $p_{\mathcal{H}}$, we find the expression for ψ_1 given in Corollary 3.4

Bibliography

Aitken, B. (1998). Have institutional investors destabilized emerging markets? *Contemporary Economic Policy*, 16(2):173–184.

Alfonsi, A., Fruth, A., and Schied, A. (2009). Optimal execution strategies in limit order books with general shape functions. *Quantitative Finance*, 10:143–157.

Allez, R. and Bouchaud, J.-P. (2012). Eigenvector dynamics: general theory and some applications. *Physical Review E*, 86(4):046202.

Almgren, R. (2009). Optimal trading in a dynamic market. *Working paper.*

Almgren, R. and Chriss, N. (2000). Optimal execution of portfolio transactions. *J. Risk*, 3:5–39.

Almgren, R. and Lorenz, J. (2006). Adaptive arrival price. *Algorithmic Trading: Precision, Control, Execution*, 3:59–66.

Almgren, R., Thum, C., Hauptmann, E., and Li, H. (2005). Direct estimation of equity market impact. *Risk*, 18(7):58–62.

Andersen, T., Bollerslev, T., Diebold, F., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71(2):579–625.

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Econometric analysis of realised covariation: high frequency covariance, regression and correlation in financial economics. *Econometrica*, 72(2):885–925.

Bence, T., Lemperiere, Y., Deremble, C., Lataillade, J. D., Kockelkoren, J., and J-P, B. (2011). Anomalous price impact and the critical nature of liquidity in financial markets. *Physical Review X*, 1(2):021006.

Bouchaud, J., Cizeau, P., Laloux, L., and Potters, M. (2000). Random matrix theory and financial correlations. *International Journal of Theoretical and Applied Finance*, 3(3):391–398.

Boudreaux, K. (2012). Discounts and premiums on closed-end mutual funds: A study in valuation. *The Journal of Finance*, 28(2):515–522.

Carhart, M. (2012). On persistence in mutual fund performance. *The Journal of finance*, 52(1):57–82.

Cont, R., Kukanov, A., and Stoikov, S. (2013). The price impact of order book events. *Journal of financial econometrics*, 12(1):47–88.

Cont, R. and Wagalath, L. (2013a). Fire sales forensics: measuring endogenous risk. *Mathematical Finance (accepted).*

Cont, R. and Wagalath, L. (2013b). Running for the exit: distressed selling and endogenous correlation in financial markets. *Mathematical Finance*, 23(4):718–741.

Da Fonseca, J., Gnoatto, A., and Grasselli, M. (2013). A flexible matrix libor model with smiles. *Journal of Economic Dynamics and Control*, 37(4):774–793.

Da Fonseca, J., Grasselli, M., and Tebaldi, C. (2008). A multifactor volatility Heston model. *Quantitative Finance*, 8(6):591–604.

Del Guercio, D. (1996). The distorting effect of the prudent-man laws on institutional equity investments. *Journal of Financial Economics*, 40(1):31–62.

Dempster, M., Evstigneev, I., and Schenk-Hoppé, K. (2003). Exponential growth of fixed-mix strategies in stationary asset markets. *Finance and Stochastics*, 7(2):263–276.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate garch models. *Journal of Business and Economic Statistics*, 20:339–350.

Ethier, S. and Kurtz, T. (1986). *Markov Processes: Characterization and Convergence.* Wiley.

Evstigneev, I. and Schenk-Hoppé, K. (2002). From rags to riches: on constant proportions investment strategies. *International Journal of Theoretical and Applied Finance*, 5(06):563–573.

Falkenstein, E. (2012). Preferences for stock characteristics as revealed by mutual fund portfolio holdings. *The Journal of Finance*, 51(1):111–135.

Ferreira, M. and Matos, P. (2008). The colors of investors' money: The role of institutional investors around the world. *Journal of Financial Economics*, 88(3):499–533.

Friedman, S. and Weisberg, H. (1981). Interpreting the first eigenvalue of a correlation matrix. *Educational and Psychological Measurement*, 41(1):11.

Fuhr, D. (2011). ETF Landscape. Industry Highlights. *Blackrock Presentation.*

Gabaix, X., Gopikrishnan, P., Plerou, V., and Stanley, H. (2006). Institutional investors and stock market volatility. *The Quarterly Journal of Economics*, 121(2):461–504.

Gastineau, G. (2010). *The exchange-traded funds manual*, volume 186. Wiley.

Gompers, P. and Metrick, A. (2001). Institutional investors and equity prices. *Quarterly Journal of Economics*, 116:229–259.

Gonnard, E., Kim, E., and Ynesta, I. (2008). Recent trends in institutional investors statistics. *Financial Market Trends*, 95:1–22.

Gouriéroux, C., Jasiak, J., and Sufana, R. (2009). The Wishart autoregressive process of multivariate stochastic volatility. *Journal of Econometrics*, 150(2):167–181.

Ikeda, N. and Watanabe, S. (1981). *Stochastic Differential Equations and Diffusion Processes.* Elsevier.

Kyle, A. and Xiong, W. (2001). Contagion as a wealth effect. *The Journal of Finance*, 56(4):1401–1440.

Lakonishok, J., Shleifer, A., and Vishny, R. (1992). The impact of institutional trading on stock prices. *Journal of financial economics*, 32(1):23–43.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.

Moro, E., Vicente, J., Moyano, L. G., Gerig, A., Farmer, J. D., Vaglica, G., Lillo, F., and Mantegna, R. N. (2009). Market impact and trading profile of hidden orders in stock markets. *Physical Review E*, 80(6):066102.

Mulvey, J. and Kim, W. (2008). Constantly rebalanced portfolios-is mean-reverting necessary. Technical report, working paper Princeton University, Princeton, NJ, USA.

Obizhaeva, A. (2011). Selection bias in liquidity estimates. *Working Paper.*

Sias, R. (1996). Volatility and the institutional investor. *Financial Analysts Journal*, pages 13–20.

Sias, R. and Starks, L. (1997). Return autocorrelation and institutional investors. *Journal of Financial economics*, 46(1):103–131.

Stelzer, R. (2010). Multivariate COGARCH$(1,1)$ processes. *Bernoulli*, 16(1):80–115.

Tonello, M. and Rabimov, S. (2010). The 2010 institutional investment report: trends in asset allocation and portfolio composition.