SCUOLA
NORMALE
SUPERIORE

Classe di Scienze
Corso di Perfezionamento in
Financial Mathematics
XXXIII ciclo

From Macro to Micro: Causal Inference, Firm Valuation and Trading Conditions

Settore Scientifico Disciplinare SECS-S/06

Candidate:

Francesco Cordoni

Advisors:

Prof. Fulvio Corsi
Prof. Stefano Marmi
Dr. Giulia Livieri

Anno Accademico 2020/2021

Ringraziamenti

Ringrazio i miei relatori Fulvio, Giulia e Stefano, che con il loro sostegno mi hanno supportato e sopportato in questi anni.

Ringrazio anche i miei coautori Fabrizio e Giulio per i loro consigli e discussioni sempre preziose.

Vorrei inoltre ringraziare tutti i professori che hanno contribuito alla mia formazione.

Ringrazio infine tutti coloro che mi sono stati vicino in questi anni. Tanti se ne sono andati e altrettanti sono rimasti.

ABSTRACT

The aim of the Ph.D. thesis is twofold. First, we investigate possible stock market mispricing to eventually build profitable investment strategies. Second, we analyze how the microscopic interactions among agents influence trading conditions thereby leading to market instabilities.

As regards the study of possible mispricing, we identify via the vector autoregressive approach revenues as the primary driver process of firm growth. To do so, we employ the recent Independent Component Analysis (ICA) technique which allows us to identify contemporaneous causal relations among the considered variables. In particular, the first original contribution of the thesis is to extend the ICA methodology for singular and noisy structural vector autoregressive models; see Chapter 2.

As a second original contribution, starting from the revenues, we propose a firm valuation framework incorporating the associated intrinsic uncertainty. We derive a probability distribution of fair values, we construct a market factor capturing misvaluation comovements and we propose two stock recommendation systems that hinge on the fair value distribution; see Chapters 3, 4 and 5.

Finally, in the last contribution, we analyze asymptotically market stability as the number of assets and traders increase. Market instability is defined as a result of oscillating equilibrium strategies of optimal execution problems in market impact games, where the dynamical equilibrium between the activity of simultaneously trading agents generates the price dynamics. One of the main results is the connection of market instability to the market cross-impact structure when portfolios execution orders are considered; see Chapter 7, 8 and 9.

TABLE OF CONTENTS

INTRODUCTION

The value of a company can be obtained by looking either at the *market* or at the *intrinsic* value. There is a significant difference between these two quantities. The former is a measure set by the market of how much it would cost to buy the company, i.e., it reflects the company market price. On the other hand, the firm's actual value can be also estimated by considering its fundamental data, e.g., cash flows, which define the company intrinsic (or fundamental) value.

If we denote the market price function at time t as the function which relates a stock in X to its market price, i.e.,

$$\begin{aligned} p_t \colon X &\to \mathbb{R}_+ \\ x &\mapsto p_t(x), \end{aligned}$$

we can define, for each t, an equivalence relation between assets by $A \sim B \iff p_t(A) = p_t(B)$. In other words, if two assets have the same market price, then they are essentially indistinguishable from each other from a market perspective. However, the intrinsic value defines another equivalence relation, $A \sim^v B \iff v_t(A) = v_t(B)$, where $v_t(\cdot)$ is the function which relates a stock with its intrinsic value. Typically, the value of publicly traded companies is related to the intrinsic value derived by fundamental data, such as income statement, balance sheet and statement of cash flows. The sticking point is that while p_t is directly observed on the market, the intrinsic value is a latent quantity. However, the underlying pragmatic assumption of financial analysts and practitioners is that the difference between the value and price will become negligible, i.e.,

$$\lim_{t\to\infty} p_t(x) - v_t(x) = 0 \quad \forall x \in X, \tag{0.1}$$

which represents a some degree of market efficiency[1]. Under this hypothesis, undervalued assets should recover, at least in part, their true value and thus, their prices are expected

[1]We adopt the point of view of Black (1986): "The noise that noise traders put into stock prices will be cumulative, in the same sense that a drunk tends to wander farther and farther from his starting point. Offsetting this, though, will be the research and actions taken by the information traders. The farther the price of a stock gets from its value, the more aggressive the information traders will become. More of them will come in, and they will take larger positions. They may even initiate mergers, leveraged buyouts, and other restructurings. Thus the price of a stock will tend to move back toward its value over time", (p. 533).

to increase[2]. Since the seminal work of Shiller (1981), many researchers have focused on how rational assumptions related to market efficiency, like rational expectation hypothesis, cannot explain the substantial difference between market price and its related fundamental value, the so called excess volatility puzzle.

One of the most fundamental direct valuation method is the discounted cash flows (DCF) approach. It is widely employed by sell-side financial analyst and practitioners (see e.g. Brown et al., 2015). In DCF valuation, one starts by determining the stream of future cash flows of a company and then computes their present value through an appropriately defined discount rate. We first have to determine the stream of future cash flows, so to reduce them to present value through the discount rate and, finally, to compare the result with today's price. This approach attempts to capture with one single number, the discount rate, two different effects: the time value of money and the uncertainty of future cash flows. Although, it is a well established fact that the market price nature of a stock follows a stochastic process, the same assumption is not widely accepted for the fundamental value in a simple framework like the DCF approach. Indeed, most of applications of the DCF are regarding the computation of the so called expected future cash flows, from which one can recover a point fair value estimate.

However, noise and uncertainty play a major role in valuation: all estimates of value are noisy, so we can never know by how much market price has deviated from value, (Black, 1986, p. 533). Noise and uncertainty limit the extent to which an analysis of cash flows can provide useful and accurate information about the value of an asset, and they both affect the determination of target prices and the bottoms-up construction of portfolios. These limitations have been evident since the early history of financial analysis: in their foundational study on security analysis, (Graham and Dodd, 1934, p. 18) state: "The essential point is that security analysis does not seek to determine exactly what is the intrinsic value of a given security. It needs only to establish either that the value is adequate - e.g., to protect a bond or to justify a stock purchase - or else that the value is considerably higher or considerably lower than the market price. For such purposes, an indefinite and approximate measure of the intrinsic value may be sufficient." Therefore, primarily due to the intrinsic difficulty of estimating the future cash flows of a company, the value provided by DCF is likely to be affected by a considerable amount of uncertainty. Starting from this consideration, existing works have highlighted the necessity of developing probabilistic and statistical tools to extend the conventional DCF method to include some measure of uncertainty associated with the estimated value (Casey, 2001). However, to the best of our knowledge, despite its practical relevance, this problem has been the subject of surprisingly few academics studies. The general

[2]Even if the fundamental value may vary on time, it is conventional wisdom that the velocity of its change is smaller than those of market price. This is also related to the frequency with the balance sheet data are released (typically quarterly).

suggestion has been to perform Monte Carlo simulations of the underlying (accounting) variables starting from historically estimated correlation matrices (see Damodaran, 2007, French and Gabrielli, 2005). This is similar to Monte Carlo procedures commonly used by analysts in investment banking studies (see e.g Koller et al., 2010). For instance, in Ali et al. (2010), Gimpelevich (2011), and Samis and Davis (2014) both scenario-based and simulation-based analyses are used, together with the DCF, for investment decisions in real-estate projects or for evaluating a specific market sector. Interestingly, in Viebig et al. (2008) the authors acknowledge that "Being intellectually honest, financial analysts can at best determine ranges or *distribution* of possible fundamental financial values but not exact price targets for stocks as future revenue growth rates, future operating margins and other inputs which go into DCF models cannot be predicted with certainty."

We propose a simple, general and theoretically grounded method, the Stochastic Discounted Cash Flow (SDCF) approach, to replace the point estimate of the conventional DCF method with a proper random variable. The basic idea of the SDCF is to consider a suitably defined probability space that is able to describe the future cash flow dynamics of a company. Should the true cash flow process be known, the value computed by the standard DCF would be exactly the expectation of the SDCF random variable. In general, the reliability of the method depends on the goodness of the data generating process describing future cash flows. Therefore, before presenting the stochastic framework to DCF approach, we have to select a well-defined process on which we may define a suitable probability space. Thus, to pursue this scope we study the causal structure of firm growth. We employ a Vector Autoregressive (VAR) model approach, a statistical model which describes the joint evolution of N observed variables with their past values. Since the seminal paper of Sims (1980), Structural Vector Autoregressive (SVAR) models are employed to study theoretical economic models and policies, such as the effect of a monetary intervention. They are one of the most prevalent tools in empirical economics to study dynamic systems of selected variables from which the causal relationships among the selected variables may be recovered. The success of the VAR approach can be attributed to its simplicity and independence from structural underlying (economic) model. Traditionally, the structural counterpart can not be identified without imposing (economic) restrictions, which are often derived by an underlying (economic) theory and thus they invalid the data-driven feature of the reduced VAR model.

However, recent works shown that this problem may be overcame without any restrictions exploiting the non-gaussianity of data via the Independent Component Analysis (ICA), Moneta et al. (2013), Lanne et al. (2017), Gouriéroux et al. (2017), a machine learning technique which allows to recover the causal properties employing only the features of the data. Even if the classical application of the SVAR approach is related to macroeconomics studies, following Coad and Rao (2010) and Moneta et al. (2013), we employ the ICA technique and a SVAR analysis to show how revenues constitute the main

driving factor of the company growth and so they may be selected as the process on which to build our stochastic valuation approach. So far, however, applications of ICA in the economic literature have been limited only to the determined case, i.e., when the number of the unobserved structural shocks R is equal to the number of observed variables N. We propose generalizations of the ICA identification technique for SVAR models in two directions: (i) to the *singular* case, i.e., when the number of variables is larger than the number of the structural shocks, $N > R$, and (ii) to the *noisy* case, i.e., when the system is contaminated with noise in the VAR residuals. The basic idea is to combine methods for the estimation of factor models together with ICA approach. In particular, we propose a combination of the collapsing procedure of Jungbacker and Koopman (2008) together with the recent ICA estimator introduced by Gouriéroux et al. (2017) allowing to achieve identification also in the singular and noisy case. We study the singular and noisy ICA formulation employing the Pseudo Maximum Likelihood (PML) estimator of Gouriéroux et al. (2017) and we show the consistency of the proposed combined procedure. A similar problem has been analyzed in the ICA literature but, in contrast to our work, these works have focused on the algorithmic performances and techniques rather than on the formal analysis of the statistical properties of the procedures, see e.g., Moulines et al. (1997), Hyvärinen (1999), Attias (1999), Joho et al. (2000). As customary with VAR analysis we apply the new identification method to U.S. macroeconomic data. Thanks to the proposed procedure we employ the entire set of information contained in all the observed macro-variables available to consistently identify the lower dimensional system of the structural shocks driving the U.S. economy.

Then, we turn to our primary task, and we propose our SDCF approach by relying on the revenues process as the main driving factor for firm growth and cash flow dynamics. In order to build a satisfactory prediction of future cash flows we rely upon two empirical observations. The first observation is that the basic source of company's cash flow, the revenues, is characterized by a substantially volatile idiosyncratic component. The second observation is that even if, from an accounting point of view, the cash flow is a reconstructed variable that depends on a set of other, more fundamental, variables (e.g. amortization, cost of debts and taxes), all interacting and affecting the final realized cash flow in different degrees, the structural relationship among these variables results stable in time.

The main methodological novelty of our approach is merging these two observations in a three steps procedure to derive a prediction model for future cash flows. First, a set of econometric models are estimated at firm level, their efficiency is individually and independently compared, and the best model of each firm is used in a Monte Carlo procedure to obtain the distribution of future revenues. Second, all the other accounting variables that enters into the final definition of company's cash flow are estimated as "margins" on the revenues by using historical data. Finally, the obtained data generating

process is used in a controlled Monte Carlo simulation to derive a probability distribution for the company's fair value, the *fair value distribution*, which may be used to obtain both an estimate of the expected fair value of the company and of its degree of uncertainty.

To explore the information content of the fair value distribution, we build a volatility-adjusted mispricing indicator, defined as the difference between the market price of the company and its expected fair value divided by the fair value distribution standard deviation. Under the assumption that company's future market prices will eventually adjust to re-absorb company's present misvaluation, we run a series of empirical exercises to investigate the relation between our misvaluation indicator and market returns. We start with a firm-level investigation. We find that the misvaluation indicator possess a significant predictive power for one-quarter ahead excess returns when used to augment linear factor model routinely used in financial applications (see e.g. the Fama-French three-factor model, Fama and French, 1993, 2015) and other control variables. To further assess the reliability of our misvaluation indicator, we sort stocks into (appropriately defined) quantiles based on the empirical distribution function of the individual firm indicator and we construct "Buy", "Hold" and "Sell" portfolios according to this quantile splitting. By comparing the equally weighted daily returns of these portfolios we observe that the "Buy" portfolio earns a gross return that is consistently and significantly higher than that of the "Sell" portfolio. Motivated by the evidence at firm level, we explore if, and to what extent, our misvaluation index has some predictive power when augmenting traditional market factor models. We form a long-short valuation factor (LSV) by measuring the returns on a factor-mimicking portfolio that goes long on the most recommended (undervalued) stocks and short in the less recommended (overvalued) stocks. In this respect, this study is related to the works of Hirshleifer and Jiang (2010) for the UMO factor and by Chang et al. (2013) for the MSV factor. The LSV factor, when added to the Fama French five-factors model (Fama and French, 2015) augmented by the momentum factor introduced in Carhart (1997), as well as by the UMO factor of Hirshleifer and Jiang (2010), is not redundant in describing average returns, both on the cross-section of portfolio and individual stock returns. This confirms the ability of our indicator of capturing a previously unexplained contribution to company's mispricing.

Then, we investigate whether investors can profit from two stock recommendation systems constructed using the information provided by the stochastic discounted cash flow analysis. The recommendation systems we introduce, based on the SCDF method, compare the fair value distributions with market prices to detect mispriced assets. The prices of undervalued assets are expected to increase. As such, they are good candidates for buy portfolios. Conversely, overvalued assets are expected to see a decrease in price and are good candidates for sell portfolios. The misvaluation assessment is run differently in our two recommendation systems. The first system is called *single-stock quantile* and the degree and direction of mispricing for a company are derived from the likelihood of

obtaining the observed market price from the fair value distribution of that company. In the second proposed system, the *cross-sectional quantile*, mispricing information is derived comparatively, jointly using the fair value distributions of all considered companies. In this approach, the mispricing indicator is the previously defined z-score index. The advantage of the former method is that it uses all information available from the fair value distribution, while the latter derives its robustness with respect to possible fair value estimation biases. The performance analysis of the two systems is carried out using an intercept test of excess portfolio returns, built using the Fama-French three-factor model (Fama and French, 1993) augmented with the momentum factor (Carhart, 1997). We show that equally weighted portfolios composed of the most highly recommended stocks consistently earn positive abnormal gross returns. The comparison with the much weaker results obtained using similar portfolios built from analyst recommendations from the Thomson Reuters Institutional Brokers' Estimate System (I/B/E/S) database further emphasizes the advantage of our methodology. Moreover, our recommendation systems remain profitable even when a high level of turnover costs is included. We believe that this study complements the literature that highlights the necessity for developing probabilistic and statistical tools to extend the conventional DCF approach by including some measure of uncertainty associated with the estimated value, Casey (2001), Bradshaw (2004), Brown et al. (2015), Baule and Wilke (2016). Finally, as a robustness check of the cross-sectional quantile methodology, we built an analogues system based on analysts recommendations and we find that it improves the original analysts' indications.

In the first part of the thesis some market efficiency hypothesis is assumed to be valid. One of the first attempts whereby market efficiency was rationalized is the introduction of the rational expectation hypothesis, where agents are assumed to be rational and perfectly informed, O'Hara (1997), Guéant (2016), Vodret et al. (2020). A particular class which incorporates these assumptions are the information-based models, where market participants are divided in two main categories, noise traders and informed traders, Kyle (1985). In particular, a noise trader is an agent which trades on the basis of exogenous (w.r.t. market) information, while an informed trader is essentially an arbitrageur, which wants to try to profit from a private information. Usually, the trading mechanisms is organized with market makers (dealers) which actually provide liquidity and make market more transparent. They display "bid" and "ask" price at which they willing to buy and sell a stock, respectively. However, both informed and noisy traders typically fragment their orders in small pieces so as to hide their private informations, e.g., using metaorders Bouchaud et al. (2009). This generates a correlation in order flows and a subsequent price impact. Market impact is the relation between orders and price changes. A popular model which describes how past orders influence the price is the *transient impact model*, where the market impact is modeled by a specific decay kernel function, also called propagator,

see for instance Bouchaud et al. (2004) Bouchaud et al. (2009) Guéant (2016). Even if we have profitable portfolios strategy, the execution in real market of large orders (like a rebalancing in a portfolio) it is not trivial, and a bad orders scheduling can deteriorate the profitability of the all investment strategy. Thus, in the second part of the thesis, we focus on the optimal schedule problem of an agent which trades many assets in a market with many competitors.

However, a non-trivial impact function is introduced to preserve the statistically efficiency of markets, which might trigger market instabilities and flash crashes, Vodret et al. (2020). Instabilities in financial markets have always attracted the attention of researchers, policy makers and practitioners in the financial industry because of the role that financial crises have on the real economy. Despite this, a clear understanding of the sources of financial instabilities is still missing, in part probably because several origins exist and they are different at different time scales. The recent automation of the trading activity has raised many concerns about market instabilities occurring at short time scales (e.g. intraday), in part because of the attention triggered by the Flash Crash of May 6th, 2010 (Kirilenko et al., 2017) and the numerous other similar intraday instabilities observed in more recent years (Brogaard et al., 2018, Calcagnile et al., 2018, Golub et al., 2012, Johnson et al., 2013), such as the Treasury bond flash crash of October 15th, 2014. The role of High Frequency Traders (HFTs), Algo Trading, and market fragmentation in causing these events has been vigorously debated, both theoretically and empirically (Brogaard et al., 2018, Golub et al., 2012).

One of the puzzling characteristics of market instabilities is that a large fraction of them appear to be endogenously generated, i.e., it is very difficult to find an exogenous event (e.g. a news) which can be considered at the origin of the instability (Cutler et al., 1989, Fair, 2002, Joulin et al., 2008). Liquidity plays a crucial role in explaining these events. Markets are, in fact, far from being perfectly elastic and any order or trade causes prices to move, which in turn leads to a cost (termed slippage) for the investor. As mentioned above, in order to minimize market impact cost, when executing a large volume it is optimal for the investor to split the order in smaller parts which are executed incrementally over the day or even across multiple days. One of the origin of the market impact cost is predatory trading (Brunnermeier and Pedersen, 2005, Carlin et al., 2007): the knowledge that a trader is purchasing progressively a certain amount of assets can be used to make profit by buying at the beginning and selling at the end of the trader's execution. Part of the core strategy of HFTs is exactly predatory trading. Now, the combined effect on price of the trading of the predator and of the prey can lead to large price oscillations and market instabilities. In any case, it is clear that the price dynamics is the result of the (dynamical) equilibrium between the activity of two or more agents simultaneously trading.

This equilibrium can be studied by modeling the above setting as a market impact

game (Carlin et al., 2007, Lachapelle et al., 2016, Moallemi et al., 2012, Schied and Zhang, 2018, Schöneborn, 2008, Strehle, 2017a,b). In a nutshell, in a market impact game, two traders want to trade the same asset in the same time interval. While trading, each agent modifies the price because of market impact, thus when two (or more) traders are simultaneously present, the optimal execution schedule of a trader should take into account the simultaneous presence of the other trader(s). As customary in these situations, the approach is to find the Nash equilibrium, which in general depends on the market impact model.

Market impact games are a perfect modeling setting to study endogenously generated market instabilities. A major step in this direction has been recently made by Schied and Zhang (2018). By using the transient impact model of Bouchaud et al. (2004, 2009) plus a quadratic temporary impact cost (which can alternatively be interpreted as a quadratic transaction cost, see below), they have recently considered a simple setting with two identical agents liquidating a single asset and derived the Nash equilibrium. Interestingly, they also derived analytically the conditions on the parameters of the impact model under which the Nash equilibrium displays huge oscillations of the trading volume and, as a consequence, of the price, thus leading to market instabilities[3]. Specifically, they proved the existence of a sharp transition between stable and unstable markets at specific values of the market impact parameters.

Although the paper of Schied and Zhang highlights a key mechanism leading to market instability, several important aspects are left unanswered. First, market instabilities rarely involve only one asset and, as observed for example during the Flash Crash, a cascade of instabilities affects very rapidly a large set of assets or the entire market (CFTC-SEC, 2010). This is due to the fact that optimal execution strategies often involve a *portfolio* of assets rather than a single one (see, e.g. Tsoukalas et al., 2019). Moreover, commonality of liquidity across assets (Chordia et al. (2000) and cross-impact effects, (Alfonsi et al., 2016, Schneider and Lillo, 2019), make the trading on one asset triggers price changes on other assets. Thus, it is natural to ask: is a large market more or less prone to market instabilities? How does the structure of cross-impact and therefore of liquidity commonality affect the market stability? A second class of open questions regards instead the market participants. Do the presence of more agents simultaneously trading one asset tends to stabilize the market? While the solution of Schied and Zhang considers only two traders, it is important to know whether having more agents is beneficial or detrimental to market stability. For example, regulators and exchanges could implement mechanisms to favor or disincentive participation during turbulent periods. Answering this question requires solving the impact game with a generic number of agents and it is discussed in

[3]In their paper, Schied and Zhang interpret the large alternations of buying and selling activity observed at instability as the "hot potato game" among HFTs empirically observed during the Flash Crash (CFTC-SEC, 2010, Kirilenko et al., 2017).

the recent work of Luo and Schied (2020). Furthermore, they also extended the original framework by considering the agents' risk aversion and the related mean-variance and CARA optimization problems. In particular, they derived explicit solutions for the corresponding Nash equilibria and they studied numerically how the stability is influenced by the presence of many agents. We extend considerably the setting of Schied and Zhang by answering the above research questions. Specifically, starting from Luo and Schied (2020), we consider (i) the case when agents trade multiple assets simultaneously and cross market impact is present and we provide explicit representations of related Nash equilibria; (ii) after studying how trading conditions may be affected by the cross impact we derive theoretical results on market stability for the $J = 2$ agents by showing how it is related to the cross-impact effect; (iii) we study numerically market stability in the general case and we extend a previous result of Luo and Schied (2020) in the multi-asset case. The different 'paths' leading to market instability are therefore highlighted, finding, surprisingly, that larger and more competitive markets are more prone to market instability. Moreover, we also exhibit a possible way to reduce these instabilities which a policy regulator would like to prevent.

The thesis is organized as follows. In Chapter 1 we quickly recall the standard DCF setting and we present the causal analysis to firm growth using SVAR-ICA approach. In Chapter 2 we extend the ICA technique in the SVAR context and we discuss under which assumptions the proposed procedure is consistent and feasible. In Chapter 3 we discuss the details of the proposed SDCF model. In Chapter 4 we introduce our misvaluation measure and we investigates the relation between stock misvaluation and future stock returns by constructing a misvaluation factor. In Chapter 5 we analyze the performance of the two proposed stocks recommendations systems. Then, in Chapter 6 we recall some notation of the market impact games framework and the Luo and Schied (2020) model. We extend the basic model of Luo and Schied (2020) to the multi-asset case in Chapter 7, where we find the corresponding Nash equilibria for different objective functions. We analyze how trading strategies are influenced by trading conditions in Chapter 8. Finally, in Chapter 9 we study market stability.

Part of the thesis content is inspired by the following research articles. In particular, Chapter 2 is referred to [1], Chapters 3 and 4 to [2], Chapter 5 to [3] and Chapters 6, 7, 8 and 9 are related to [4].

[1] F. Cordoni and F. Corsi. Identification of Singular and Noisy Structural VAR Models: The Collapsing-ICA Approach.
Available at SSRN: https://ssrn.com/abstract=3415426, 2019.

[2] G. Bottazzi, F. Cordoni, G. Livieri, and S. Marmi. Uncertainty in Firm Valuation and a Cross-Sectional Misvaluation Measure.
Available at SSRN: https://ssrn.com/abstract=3614988, 2020.

[3] G. Bottazzi, F. Cordoni, G. Livieri, and S. Marmi. Stock Recommendations from Stochastic Discounted Cash Flows.
Available at SSRN: https://ssrn.com/abstract=3621514, 2020.

[4] F. Cordoni and F. Lillo. Instabilities in Multi-Asset and Multi-Agent Market Impact Games.
arXiv preprint arXiv:2011.10242, 2020.

Part I

Causal Inference and Firm Valuation

CHAPTER 1

The Valuation of a Firm

Valuation is the process from which the fair value of a company is recovered. The price of an asset refers to the consensus of agents about the amount of money that they are willing to pay to buy that asset. On the other hand, in the "valuation process" of a firm we are interested to determine the value of a company characterized by its intrinsic properties. The literature on company valuation encompasses three estimation approaches of the shareholder value. The first "direct valuation" approach estimates the value of a firm directly from its expected future cash flows, without any reference to the current price of other firms in the market. Common examples of this approach are the dividend discount model, the residual income model and the discounted cash flow model. The second "relative valuation" approach estimates firm value by examining the market pricing of assets comparable to those composing the company under valuation, typically in relation to some accounting variable. The third, less common, approach is constituted by the "contingent claim valuation", which derives company's value starting from the prices of related options (see Damodaran (2012) for further details). In this Chapter we are going to review some techniques related to the standard DCF model and we identify, via structural analysis, the main driver process of firm growth which is at the basis of the proposed stochastic DCF model in Chapter 3.

1.1 The Discounted Cash Flow model

In the DCF model the firm fair value is defined by discounting expected future cash flows. Therefore, the definition of cash flows is a crucial step for the DCF method. Essentially, the cash flows should represent the amount of excess cash generated by a company, which may be used to invest for the business and enrich shareholders. For this reason, in order to distinguish them to the standard operating cash flows[1], they are often referred to as *free cash flows*. Broadly speaking, they are obtained by the difference of the capital expenditures to the net cash flows from operating activities and they constitute a measure of liquidity. For the sake of simplicity we denote by CF_t the free cash flow of a company at time t. The DCF method is founded in the present value approach and the

[1] Which are a measure of the amount of cash generated by standard business operations of a company.

value of a firm, v_0, is essentially given by the present value of the expected future cash flows appropriately discounted,

$$v_0 = \sum_{t=1}^{\infty} \frac{\mathbb{E}[CF_t]}{\prod_{s=1}^{t}(1+k_s)}, \tag{1.1}$$

where $\mathbb{E}[\cdot]$ denotes the expected value operator and k_s, $1 \leq s \leq t$, denotes the appropriate discount factor at time s. The main difference among various DCF models are regarding the growth and the definition of cash flows.

If we assume that the cash flows have a constant growth rate, then we obtain the so called one-stage model, where the value of a firm is essentially given by expected cash flow in the next period over the difference between the discount rate and the stable growth rate of cash flows, g, i.e.,

$$v_0 = \frac{\mathbb{E}[CF_1]}{k-g},$$

which is related to the famous Gordon growth model, Gordon (1962). In a two-stage model it is assumed a faster firm growth in an initial period subsequently followed by a stable phase, while in the three-stage model, there are an initial phase of high growth, a transitional period in which the growth rate tend towards a steady-state phase in which the firm growth is stable. The choice of the growth rates, which are typically selected from historical data, is a crucial step for a DCF model, which turns out to be very sensitive to a small variation of these rates. However, we discuss in Chapter 3 how we can overcome these difficulties when, once identified a suitable driver process for cash flows dynamics, we select a suitable structural model for this kind of process. Even if there is no a universal model for company valuation, we will focus on the two-stage framework which represents a good trade-off between the simple Gordon and the three-stage models[2]. Then it is assumed, see e.g., Damodaran (2007), Steiger (2010), that

Assumption 1. The discount factor is constant in time and it is denoted by k.

Assumption 2. There exist a date $T > 0$ and a rate g, where $0 < g < k$ and k as in Assumption 1, called "perpetual growth rate" such that

$$CF_{t+1} = CF_t \cdot (1+g), \ \forall\, t \geq T.$$

Under Assumption 1 and 2 the infinite summation in Eq. (1.1) can be rewritten as:

$$v_0 = \sum_{t=1}^{T} \frac{\mathbb{E}\left[CF_t\right]}{(1+k)^t} + \frac{\mathbb{E}\left[CF_T\right](1+g)}{(1+k)^T(k-g)}, \tag{1.2}$$

where the second term multiplied by $(1+k)^T$ is the so called *terminal value* (TV).

[2]Moreover, it appears to be one of the most employed model by analysts and it is the standard procedure adopted also by our data provider, Eikon/Datastream.

Another classification between DCF models is due the definition of free cash flows. In DCF, the cash flow is distinguished between *Unlevered Free Cash Flow* (UFCF) and *Levered Free Cash Flow* (LCFC)[3]. The main difference between the two are expenses. Precisely, UFCF represents the cash flow available to the debt and equity holders after all expenses are paid off and reinvestments are made; LFCF, instead, represents the cash flow employed to pay out dividend and financial expansion investments. Specifically, for each time t the UFCF is define as the difference between the sum of Net Operating Profit After Tax ($NOPAT$) and Depreciation and Amortization ($D\&A$) and the sum of Capital Expenditure ($CAPEX$) and Change in the Working Capital (WC), (see, e.g., Damodaran, 2007, 2012, Steiger, 2010):

$$UFCF_t := NOPAT_t + D\&A_t - CAPEX_t - \Delta WC_t, \tag{1.3}$$

where $\Delta WC_t = WC_t - WC_{t-1}$, $NOPAT_t = (EBITDA - D\&A) \cdot (1 - \tau)$, $EBITDA$ represents Earning Before Interest, Taxes, Depreciation, and Amortization and τ represents the marginal tax rate. A detailed description of these accounting variables is reported in Table 1.1. Then, the LFCF is defined by adding the net debt issued to the UFCF.

Therefore, by discounting the UFCF we can recover the firm enterprise value, while the DCF model with LCFC provides directly the equity value. Even if, we are interested in the equity valuation of a firm, we will use the UFCF as a proxy for cash flows, since their definition is less prone to arbitrary decisions, see Damodaran (2007, 2012), Steiger (2010), and moreover the UCFC DCF model is in line with the one used by our data provider Eikon/Datastream as a standard industry valuation method. Therefore, in a UFCF DCF model once obtained the enterprise value v_0 we have to remove, i.e., subtract, the current value of "debt". More precisely, the equity fair value v_0^{Eq} is defined as:

$$v_0^{Eq} = v_0 - (TD - CsI + MI + PS), \tag{1.4}$$

where TD represents the Total Debt, used as a proxy for the market value of debt (a proxy suggested also by Damodaran (2007) and Steiger (2010)), CsI the Cash and Short-term Investments, MI the Minority Interest and PS denotes the Preferred Stock; The accounting variables used in the UFCF DCF model are all described in Table 1.1.

What we are going to present in Chapter 3 is an extended valuation framework which turns out to be very flexible to the particular selections that analysts can make.

[3]The former is also known as Free Cash Flow to Firm and the latter as Free Cash Flow to Equity.

Table 1.1 Description of some accounting variables employed in the UFCF DCF method.

Variable	Description
EBIT	Earning Before Interest and Taxes (known also as Operating profit) which represents total revenues from all of a company operating activities, after deducting any sales adjustments, excise taxes and other expenses incurred from operating activities.
D&A	Depreciation and Amortization which refers to the amount of expense charged by a company to cancel the cost of a plant or machinery over its useful life, giving consideration to wear and tear, obsolescence and salvage value together with a non-cash expense incurred due to the amortization of intangible fixed assets and goodwill.
EBITDA	Earning Before Interest, Taxes, Depreciation, and Amortization.
Marginal Tax Rate	It represents the amount of tax which we have to deduct from earnings. For sake of simplicity, it is denoted by τ.
NOPAT	Net Operating Profit After Tax, which is defined, at each time instant, as the difference between $EBITDA$ and $D\&A$, multiplied by $(1-\tau)$.
WC	Working Capital which is defined as the difference between current assets and current liabilities of a company so that it represents a measure of liquidity and efficiency. We will denote by $\Delta WC_t = WC_t - WC_{t-1}$ the Change in the WC.
CAPEX	Capital Expenditure which includes all expenditures that are not treated as an expense on the income statement when they are incurred, such as the costs for the development of software which are capitalized at the time of development, and amortized later, when such software is actually implemented for production.
TD	Total Debt Outstanding which represents the total interest-bearing debt outstanding.
CsI	Cash and Short-term Investments which represents short-term, highly liquid investments that are both readily convertible to known amounts of cash together with cash on hand and any investments in debt and equity securities with maturity of one year or less.
MI	Minority Interest which represents accumulated interest for minority shareholders in subsidiaries that are less than 100 percent owned by the reporting parent company.
PS	Preferred Stock which represents the value of preferred shares.

1.1.1 The discounting rate

Let us briefly discussing some issues regarding the discounting rates of a DCF model. The choice is closely related to the type of cash flows employed. However, the main idea

is founded in actualization principles, which are the elementary financial mathematics notions related to the concept of time value. For instance, if at a future time t we are going to receive a certain quantity of money, then to evaluate that money today we have to discount it using the riskless interest rate r. However, future cash flows of a firm are not a certain quantity, so a widespread technique is to raise the riskless rate r with a risk premium z, so that we also discount for the intrinsic uncertainty inside the cash flows. Therefore, the choice of the discounting rate is related to the intrinsic uncertainty incorporated in the expected future cash flows which are discounted to to present, see Kruschwitz and Löffler (2006) for a complete dissertation. We will discuss this issue on Section 3.1.3 after introducing the stochastic discounted cash flow model. The appropriate discount rates for the LFCF is constituted by the *Cost of Equity* (k_e), which is often derived under the *Capital Asset Pricing Model*, see Kruschwitz and Löffler (2006), Steiger (2010), Damodaran (2012), so that $k_e := r_f + \beta(r_M - r_f)$, where r_f is the risk free rate, β incorporates the risk of holding the firm stock that must be added to the investor's portfolio and r_M is the market return. An usual procedure is to replace β with the so called *levered beta*, $\beta_L := \beta[1 + (1-\tau)(Debt/Equity)]$, where τ is the marginal tax rate and *Equity* is market value of equity.

If UFCFs are employed then we have to combine the risk related to equity holder with that of debt. Then, the resulting discount rate is the *Weighted Average Cost of Capital* (WACC),

$$k = k_e w_e + k_d w_d + k_p w_p, \quad w_e + w_d + w_p = 1,$$

where k_d represents the cost of debt, k_p is the costs of preferred stocks and w_e, w_d, w_p are the related weights[4]. The cost of debt measures the current cost to the firm to pay on its outstanding debt. A common procedure is to compute the pre-tax cost of debt $k_{pre,d}$ by adding the riskless rate to the default spread, obtained using the company's credit rating and since interest rate costs are tax deductable, $k_d := k_{pre,d} \cdot (1 - \tau_{eff})$ where τ_{eff} represents the effective tax rate payed by the company. The preferred stocks constitute an hybrid financing instrument which presents a combination of equity and debt features. Their associated costs can be computed, see Damodaran (2012), as the ratio of the preferred dividend per share over the market price per preferred stocks. For the empirical application illustrated in the next chapters, see Chapters 3, 4 and 5, these values are provided directly from Thomson Reuters Eikon, Datastream database and we refer to Damodaran (2007, 2012) for a detailed description of their computation. Finally, as usual, the proxy of the perpetual growth rate g is the riskless T-bond rate.

[4]e.g., $w_e = Equity/Assets$, $w_d = Debt/Assets$, and $w_p = Preferred\ Stocks/Assets$.

1.2 What is the driver process behind firm growth?

In order to select an appropriate driver process for cash flows, we test the conventional wisdom to consider the revenues as the driver process of firm growth. Even if, it seems unnecessary to test this claim, we analyze the causal structure of accounting variables employing a SVAR technique. Moneta et al. (2013) show that sales constitute the driving factor for firm growth. In this section, we replicated their study on our data. Then, following their analysis and Coad and Rao (2010), we first select one sector from our universe of S&P500 stocks, i.e., Technology Firms (ICB 9000), which is the sector with the greatest average market capitalization in our selected universe of stocks[5]. Thus, we first recall the general methodology of SVAR analysis and its related identification problem.

1.2.1 The SVAR framework

A VAR model is a statistical model who describes the joint evolution of multivariate time series with their past values. In particular, a VAR(p) model is given by the following equation:

$$\boldsymbol{y}_t = \boldsymbol{v} + A_1\boldsymbol{y}_{t-1} + \cdots + A_p\boldsymbol{y}_{t-p} + \boldsymbol{u}_t, \tag{1.5}$$

where $\boldsymbol{y}_t \in \mathbb{R}^N$ is the vector of the observable time series of interest, $\boldsymbol{v} \in \mathbb{R}^N$ is the intercept, $A_1, A_2, \ldots, A_p \in \mathbb{R}^{N\times N}$ are parameter matrices and $\boldsymbol{u}_t \in \mathbb{R}^N$ are the residuals which are serially uncorrelated with mean zero and covariance matrix $\Sigma_{\boldsymbol{u}}$. We denote with $\Sigma_{\boldsymbol{x}}$ the variance-covariance matrix of a vector $\boldsymbol{x}_t$. The previous form is also named *reduced form* of a VAR(p) model, which is actually what one can estimate from the empirical data and it is widely used to forecast the selected variables. Then, vector $\boldsymbol{u}_t$ corresponds to the so called *reduced shocks*. Since the seminal paper of Sims (1980), the VAR approach is employed to analyze causal properties with the so named *structural* representation. Therefore, let $\boldsymbol{\alpha}_t \in \mathbb{R}^N$ be the vector of *structural shocks*, which are temporally uncorrelated, strictly stationary, with zero mean, uncorrelated $\mathrm{Cov}(\boldsymbol{\alpha}_t) = \Sigma_{\boldsymbol{\alpha}} = I_N$ and such that $\boldsymbol{u}_t = Z\boldsymbol{\alpha}_t$, where $Z \in \mathbb{R}^{N\times N}$ is nonsingular and it is called *mixing matrix* which satisfies $\Sigma_{\boldsymbol{u}} = ZZ'$. In other words, the observable reduced shocks are a mixture of the respective structural shocks, which characterize the structural VAR model. Once we have estimated the reduced form, if we multiply Eq. (1.5) by the inverse of Z we obtain the structural VAR formulation, $SVAR(p)$:

$$A_0\boldsymbol{y}_t = \boldsymbol{v}^* + A_1^*\boldsymbol{y}_{t-1} + \cdots + A_p^*\boldsymbol{y}_{t-p} + \boldsymbol{\alpha}_t, \tag{1.6}$$

[5]We postpone to Section 3.2 the accurate description of our universe of stocks, after presenting the employed econometric models for the empirical analysis of the SDCF method.

where $A_0 = Z^{-1}$, $\boldsymbol{v}^* = A_0\boldsymbol{v}$ and $A_j^* = A_0 A_j$, $j = 1, 2, \ldots, p$. Using this representation it is possible to recover the causal dependencies among each variable. However, without additional restrictions Z cannot be identified, since we can trivially replace Z by ZC, where C is an orthogonal matrix, and obtain a different structural representation starting from the same reduced form. Indeed, from Eq. (1.5) we can estimate the covariance matrix $\boldsymbol{\Sigma_u}$ and since we are looking for a matrix such that $\boldsymbol{u}_t = Z\boldsymbol{\alpha}_t$ where $\boldsymbol{\Sigma_\alpha} = I_N$, for any orthogonal matrix C,

$$\boldsymbol{u}_t = \underbrace{ZC}_{\widetilde{Z}}\underbrace{C'\boldsymbol{\alpha}_t}_{\widetilde{\boldsymbol{\alpha}_t}},$$

so that we can obtain another structural representation, which is driven by different structural shocks $\widetilde{\boldsymbol{\alpha}_t}$. This problem is termed *identification problem* of SVAR model, and it arises since the reduced form does not provide enough observations to identify the mixing matrix Z, e.g., the observed covariance matrix $\boldsymbol{\Sigma_u}$ provided only $N \cdot (N+1)/2$ conditions, since it is symmetric, while Z has N^2 unknowns. Therefore, typical approaches are to impose restrictions on the mixing matrix in order to obtain the same number of conditions, e.g., short-run (as Cholesky), long-run, signs restrictions, which are often derived by an underlying economic theory. However, a characteristic which contributes to the success of the VAR approach is its independence from structural economic model and imposing economic restrictions may invalid the data-driven feature of the VAR model. On the other hand, more recent works have shown that it is possible to identify the structural form without imposing any type of economic restrictions, exploiting either heteroskedasticity, see e.g. Sentana and Fiorentini (2001), Rigobon (2003), Lütkepohl and Netšunajev (2017), or non-gaussianity via the Independent Component Analysis approach, see e.g. Moneta et al. (2013), Lanne et al. (2017), Gouriéroux et al. (2017), Gouriéroux et al. (2019), Gagliardini and Gouriéroux (2019). We will follow the ICA approach and we recall this general methodology in the next section. As usual, only stationary time series are considered, i.e.,

$$\det A(z) := \det(I_N - A_1 z - \cdots - A_p z^p) \neq 0, \ |z| \leq 1, \ z \in \mathbb{C}.$$

1.2.2 The Independent Component Analysis

The ICA is part of a rich machine learning literature concerning the more general *Blind Source Separation* (BSS) Problem, defined as follows. Let consider N observed variables $\boldsymbol{u} \in \mathbb{R}^N$ which are linear combination of N unobserved sources $\boldsymbol{\alpha} \in \mathbb{R}^N$:

$$\boldsymbol{u} = Z\boldsymbol{\alpha}, \quad Z \in \mathbb{R}^{N\times N},$$

where $Z \in \mathbb{R}^{N\times N}$ $\det(Z) \neq 0$. As usual in this context, $\boldsymbol{u}$ are called *sensors*, Z is called *mixing matrix* and $\boldsymbol{\alpha}$ are called *sources*. Starting from a set of mixed signals, the aim

of BSS problem is to isolate the source signals without the aid of information about the mixing process.

One of the technique used for solving the BSS problem is ICA. Given a large number of T observations $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_T$ of $\boldsymbol{u}$, ICA identifies Z and (the distribution) of $\boldsymbol{\alpha}$, where

$$\boldsymbol{u}_t = Z\boldsymbol{\alpha}_t, \quad \forall t = 1, \ldots T, \quad \boldsymbol{u}_t \in \mathbb{R}^N, \quad \boldsymbol{\alpha}_t \in \mathbb{R}^N, \quad Z \in \mathbb{R}^{N \times N}.$$

The identification is guaranteed by the following Theorem of Eriksson and Koivunen (2004), (see also Theorem 11 of Comon, 1994), which is based on nice probabilistic results on the characterization of Gaussian random variables, see Darmois (1953), Skitovitch (1953), Comon (1994), Cramér (2004), Gouriéroux et al. (2017) for further details, which we report for completeness[6].

Lemma 1.1 (Cramér)**.** *If $X_1, X_2, \ldots, X_N$ are independent random variables and if $Y = \sum_{i=1}^N a_i X_i$ is Gaussian, then all the random variables X_i for which $a_i \neq 0$ are Gaussian.*

Theorem 1.2 (Darmois (1953),Skitovitch (1953))**.** *If $X_1, X_2, \ldots, X_N$ are independent random variables, let us define the two random variables Y_1 and Y_2 as*

$$Y_1 = \sum_{i=1}^N a_i X_i, \quad Y_2 = \sum_{i=1}^N b_i X_i,$$

where $a_i \neq 0$ and $b_i \neq 0$ for $i = 1, 2, \ldots, N$. Then, if Y_1 and Y_2 are independent, all variables X_i are Gaussian.

Therefore, Y_1 and Y_2 can not be independent as soon as one of the X_i is non-Gaussian.

Theorem 1.3 (Comon (1994) Th. 11, Eriksson and Koivunen (2004) Th. 3)**.** *Let $\boldsymbol{u}, \boldsymbol{\alpha} \in \mathbb{R}^N$ and $Z \in \mathbb{R}^{N \times N}$ and consider the model $\boldsymbol{u} = Z\boldsymbol{\alpha}$, where:*

- *Z is invertible;*
- *$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_N$ are independent, with at most one Gaussian distribution.*

Then, the matrix Z is identifiable up to the post multiplication by DP, where D is a diagonal matrix with non-zero elements and P is a permutation matrix.

A widely employed preprocessing step is the so called whitening process, which is used before applying ICA.

Remark 1.4 (Whitening). Given $\boldsymbol{u}_t = Z\boldsymbol{\alpha}_t$, we observe that $\Sigma_{\boldsymbol{u}} = Z\Sigma_{\boldsymbol{\alpha}}Z'$. Then, by computing a Cholesky decomposition of $\Sigma_{\boldsymbol{u}} = SS'$ we can consider the standardized sensors ($\widetilde{\boldsymbol{u}_t}$) and sources ($\widetilde{\boldsymbol{\alpha}_t}$) such that: $\boldsymbol{u}_t = S\widetilde{\boldsymbol{u}_t}$; $\boldsymbol{\alpha}_t = \Sigma_{\boldsymbol{\alpha}}^{1/2}\widetilde{\boldsymbol{\alpha}_t}$. Thus, by construction,

$$\Sigma_{\widetilde{\boldsymbol{u}}} = (S^{-1}Z\Sigma_{\boldsymbol{\alpha}}^{1/2}) \cdot (S^{-1}Z\Sigma_{\boldsymbol{\alpha}}^{1/2})' = I_N, \quad \Sigma_{\widetilde{\boldsymbol{\alpha}}} = I_N,$$

[6]For consistency with standard literature notation, we denote random variables with capital letters in Lemma 1.1 and Theorem 1.2.

and if we denote with $\widetilde{Z} = S^{-1}Z\Sigma_{\boldsymbol{\alpha}}^{1/2}$, we can reformulate the BSS problem as

$$\widetilde{\boldsymbol{u}}_t = \widetilde{Z}\widetilde{\boldsymbol{\alpha}}_t, \quad \widetilde{Z}\widetilde{Z}' = \widetilde{Z}'\widetilde{Z} = I_N.$$

Actually, a ICA problem is often formulated directly with the standardized sources and an orthogonal mixing matrix. So, if not clearly specified and for the sake of simplicity, we denote in the following with $\boldsymbol{\alpha}_t$ the standardized sources.

The whitening is essentially a pre-processing step and it is helpful for simplifying the ICA problem since the observed mixture variables are sphered with uncorrelated components and unit variance, i.e., they are whitened, see Hyvärinen et al. (2001) for further details.

The assumptions of independent sources is quite natural on BSS, while the assumptions of non-Gaussian distribution seems to be complicated and unnatural. So, why are Gaussian variables forbidden? Let us consider a simple example. If we assume that the joint distribution of two sources, α_1 and α_2, is Gaussian,

$$p(\alpha_1, \alpha_2) = \frac{1}{2\pi} \exp\left(-\frac{||\boldsymbol{\alpha}||^2}{2}\right),$$

then, the joint density of the mixture $\boldsymbol{u} = Z\boldsymbol{\alpha}$, where Z is orthogonal, is given by:

$$p(u_1, u_2) = \frac{1}{2\pi} \exp\left(-\frac{||Z'\boldsymbol{u}||^2}{2}\right) |\det(Z')| = \frac{1}{2\pi} \exp\left(-\frac{||\boldsymbol{u}||^2}{2}\right).$$

In other words, the orthogonal mixing matrix does not modify the density function and we have no further information about the independence sources, i.e., the original and mixed signals are identical. Therefore, since the spherical symmetry of the Gaussian distribution all we can do with Gaussian sources is actually whitening, see Hyvärinen et al. (2001) and for instance the illustrative example of Moneta et al. (2013). What happens if we try to use ICA and some of the components are Gaussian and some non-Gaussian? In this case we can estimate all the non-Gaussian components without separating the Gaussian ones, i.e., we can identify only the space of Gaussian sources. In other words, some of the estimated components will be arbitrary linear combinations of the original sources.

There are many numerical methods for ICA, e.g., one of the most famous is the Fast-ICA of Hyvärinen (1999). Recently method which used the Joint Approximation Diagonalization of Eigen-matrices Cardoso (1999), Bonhomme and Robin (2009) are also widely employed. For the VAR identification problem one of the most used is the VAR-LINGAM, see Shimizu et al. (2006) and Moneta et al. (2013), however, we focus on the more recent PML approach of Gouriéroux et al. (2017) which we briefly recall in the following section. We show how to extent this kind of approach to the more general problem of singular and noisy VAR model in Chapter 2.

1.2.3 ICA approach to SVAR: the PML estimator

As pointed out by Moneta et al. (2013), Lanne et al. (2017), Gouriéroux et al. (2017) the SVAR identification can be formulated as a BSS problem, where the reduced shocks are interpreted as the sensors, which are the mixed signals of unobserved sources, i.e., the structural shocks. More precisely, starting from the reduced form a VAR model (1.5), the mixed signal is given by the reduced shocks, $\boldsymbol{u}_t$ and the sources are the structural shocks $\boldsymbol{\alpha}_t$ and we have to identify the mixing matrix Z such that:

$$\boldsymbol{u}_t = Z\boldsymbol{\alpha}_t.$$

As remarked before, a whitening preprocessing is applied to the sensors, so that the problem can be reformulated as $\boldsymbol{u} = SC\boldsymbol{\alpha}$, where $C \in \mathbb{R}^{N\times N}$ is orthogonal and $S \in \mathbb{R}^{N\times N}$ is such that $\Sigma_{\boldsymbol{u}} = SS'$, so that $\Sigma_{\boldsymbol{\alpha}} = I_N$. In other words, we decompose the mixing matrix Z as the product of S and C. The matrix S can be directly estimated by $\Sigma_{\boldsymbol{u}}$, e.g., using the Cholesky decomposition. Thus, the actual problem is to estimate the orthogonal matrix C. The main assumptions for the PML approach are:

Assumptions 1.1.

i) the error process $\boldsymbol{\alpha}_t$ is i.i.d. with $\mathbb{E}[\boldsymbol{\alpha}_t] = \mathbf{0}$, $Var[\boldsymbol{\alpha}_t] = I_N$ and

ii) the components of $\boldsymbol{\alpha}_t = (\alpha_{1,t}, \alpha_{2,t}, \ldots, \alpha_{N,t})$ are mutually independent with at most one Gaussian (marginal) distribution.

In particular the true probability density functions (p.d.f.) $f_i(\alpha_i)$ of the latent components of $\boldsymbol{\alpha}_t$ are unknowns. So, Gouriéroux et al. (2017) introduce a set of p.d.f $g_i(\alpha_i)$, which are not necessarily the true ones, and they consider the pseudo log-likelihood estimator:

$$\widehat{C}_T = \arg\max_{Z} \sum_{t=1}^{T}\sum_{i=1}^{N} \log g_i(\boldsymbol{c}_i'\boldsymbol{u}_t), \quad \text{s.t. } CC' = I_N,$$

where $\boldsymbol{c}_i'$ correspond to the i-th column of C. Then, under other technical assumptions which are omitted for the sake of simplicity[7], it is possible to identify the matrix C without using any type of restrictions[8]. Moreover, it is possible to fully characterize the asymptotic properties of the PML estimator $\widehat{C}_T$.

Theorem 1.5 (Consistency of PML estimator Gouriéroux et al. (2017))**.** *Under assumptions A.1-A.5 of Gouriéroux et al. (2017), the PML estimator of C exists asymptotically and is a consistent estimator of C.*

[7] See Gouriéroux et al. (2017) for a complete discussion.

[8] For a possible way to relax assumption i) see Appendix A of Lanne et al. (2017).

Theorem 1.6 (Asymptotic Distribution of PML estimator Gouriéroux et al. (2017))**.** *Under assumptions A.1-A.5 of Gouriéroux et al. (2017) the PML estimator* $\widehat{C}_T$ *of* Z *is asymptotically normal with speed of convergence* $1/\sqrt{T}$*. The asymptotic variance-covariance of* $vec(\sqrt{T}(\widehat{C}_T - C)$ *is of the form* $A^{-1}\begin{pmatrix}\Omega & 0\\ 0 & 0\end{pmatrix}(A')^{-1}$ *where* $A \in \mathbb{R}^{N^2\times N^2}$ *and* $\Omega \in \mathbb{R}^{N\cdot(N-1)/2\times N\cdot(N-1)/2}$*.*

See Appendix B of Gouriéroux et al. (2017) for further details. Using the asymptotic distribution of the PML estimator we may test standard economic restrictions as explained by Gouriéroux et al. (2017). The matrix C can be identified up to permutations and scalings. As stressed by Gouriéroux et al. (2017) the possibility of positive scaling can be avoided by assuming the orthogonality of C, so that only permutations and change of signs are allowed. In other words, the assumption of $CC' = I$ is made also to avoid a local identification problem, i.e., the possibility of replacing C with CD where D is a diagonal matrix with strictly positive diagonal element. From the point of view of impulse responses functions analysis, which we discuss in Chapter 2, this means that only a problem of labelling and change of sings still remains after the estimation of C. However, although the identification is up to a permutation and change of signs, the economic interpretation of the structural shocks $\alpha_{j,t}$ can be based on the shapes of the $\widehat{IRF}_{i,j}(k)$ for each variables, (Gouriéroux et al., 2017). Also, it always possible to rename $-\alpha_{j,t}$ as $\alpha_{j,t}$ and to change the sign of the j-th column of Z accordingly. Another crucial feature of the PML estimator is the robustness to the misspecification of p.d.f., which has no effect on the consistency.

Remark 1.7. Another way to reformulate the BSS problem is to consider the equivalence relation w.r.t. permutations and scalings, $\sim$, in the class of nonsingular matrix. Then, ICA identifies C in an equivalence class, i.e., ICA estimates the equivalence class of C. However, it is possible to select a particular representative by using the Identification Scheme of Lanne et al. (2017), so that if $A \sim B$, then $ID(A) = ID(B)$, where ID denotes the identification scheme operator, see Lanne et al. (2017) for further details. Therefore, given two representatives A and B of two equivalence classes we want to verify whether the two equivalence classes are the same, i.e., $ID(A) = ID(B)$. For this purpose the ID may be applied on A and B.

In Chapter 2, we focus on extensions of the BSS problem in the SVAR context and in particular to the PML estimator of Gouriéroux et al. (2017). As already mentioned at the beginning of this section we now exhibit a SVAR exercise to causal identification of the firm growth process using the ICA approach, so that identification of the mixing matrix can be obtained without using any type of priori restrictions on the causal structure.

1.2.4 A causality exercise for firm growth

We are looking at the structural relations among the growth rates of revenues (REV), earning before interest, taxes, depreciation and amortization ($EBITDA$), depreciation and amortization ($D\&A$) and capital expenditure ($CAPEX$)[9]. Following Coad and Rao (2010) and Moneta et al. (2013) we restrict to a specific sector, Technology following Industry Classification Benchmark (ICB 9000), of our data sample employed successively in subsequent chapters[10]. The selected data sample starts from FQ1 1991 and ends to FQ1 2009 with a cross-section of 16 firms with a total of 1102 number of observations. For each variables we compute the growth rates[11] and we pooled all firms together in a panel VAR regression under the standard assumptions that growth process of different companies display similar structural patterns, see Coad and Rao (2010) and Moneta et al. (2013). Growth rates tend to exhibit a low degree of persistence in contrast to the dynamics of firm size levels, which motivates why a VAR model should be used for growth rates and why structural models, like the local linear trend model, are used to describes the revenues process and cash flows dynamics, see Moneta et al. (2013) and reference therein. We estimate a 1-lag[12] panel VAR model, i.e., if $\boldsymbol{y}_{i,t} \in \mathbb{R}^4$ is the vector which contains the growth rate series for the i-th firm at time t,

$$\boldsymbol{y}_{i,t} = A_1 \boldsymbol{y}_{i,t-1} + \boldsymbol{u}_{i,t}, \tag{1.7}$$

where $A_1 \in \mathbb{R}^{4\times 4}$ and $\boldsymbol{u}_{i,t} \in \mathbb{R}^4$ is the vector of the reduced shocks, or simply residuals. The growth rate series are positive correlated with an exception of the correlation between (the growth rate of) $EBITDA$ and $D\&A$. As usual, we consider the structural counterpart of Eq. 1.7 and the ICA technique to identify the mixing matrix Z such that $\boldsymbol{u}_{i,t} = Z\boldsymbol{\alpha}_{i,t}$, where $\boldsymbol{\alpha}_{i,t}$ are the structural shocks related to the respective variables. Therefore, the j-th row in the matrix Z correspond to the contemporaneous effect of a unitary shocks of the j-th variable on the others. We assume that the causal structure among the growth rates is invariant with respect to the firm. We test the Gaussianity of the residuals using the Jarque-Bera test, which reject the Gaussianity assumptions for all residuals. Although the assumptions of independence of the reduced shocks can not be tested, we will consider it appropriate, see also Moneta et al. (2013). However, the residuals do not exhibit particular significantly dependence even temporally. Therefore, the ICA approach can be employed without any particular objection. So, consistently with the previous notation, Z is decomposed as the product of the Cholesky matrix S,

[9] These are the fundamental variables of $\overline{CF}$ definition, see Section 3.1.1.

[10] See Section 3.2 for full details about data sample.

[11] Since some variables, e.g., EBITDA, exhibit both negative and positive values we use the simple growth rates.

[12] This choice is motivated since all the growth rates refer to trailing twelve months data and so it may refers to an annual base variation.

Table 1.2 Coefficients of the estimated matrix C with their relative asymptotic standard deviations in parenthesis. The coefficients in bold are significantly at significance level 1%.

	Rev.gr	EBITDA.gr	DA.gr	CAPEX.gr
Rev.gr	**0.9933**	**0.0591**	**-0.0967**	-0.0180
	(0.0015)	(0.0151)	(0.0109)	(0.0220)
EBITDA.gr	**-0.0610**	**0.9979**	**-0.0230**	0.0088
	(0.0151)	(0.0009)	(0.0072)	(0.0111)
DA.gr	**0.0953**	**0.0213**	**0.9951**	-0.0043
	(0.0109)	(0.0072)	(0.0011)	(0.0092)
CAPEX.gr	0.0263	-0.0074	0.0024	**0.9997**
	(0.0220)	(0.0111)	(0.0091)	(0.0004)

i.e., S is such that the variance-covariance of the reduced shocks is given by SS', and an orthogonal matrix C. We report the estimated C matrix, using the PML estimator of Gouriéroux et al. (2017) in Table 1.2. The matrix Z is rescaled such that the diagonal elements of $A_0 = Z^{-1}$ are equal to one. Then, we compute the mixing matrix Z, from which we recover the structural formulation[13] of Eq. 1.7,

$$A_0 \boldsymbol{y}_{i,t} = A_0 A_1 \boldsymbol{y}_{i,t-1} + \boldsymbol{\alpha}_{i,t}.$$

and the matrix of contemporaneous effect B, i.e., $A_0 = I - B$, so that

$$\boldsymbol{y}_{i,t} = B\boldsymbol{y}_{i,t} + A_1^* \boldsymbol{y}_{i,t-1} + \boldsymbol{\alpha}_{i,t},$$

where the diagonal elements of B are zero and $A_1^* = A_0 A_1$.

Table 1.3 and 1.4 display the elements of matrix B and A_1^*, respectively, where we have also computed the related standard errors using 100 bootstrap samples on residuals.

From the elements of the contemporaneous effect matrix B, it is clear how there is no significant contemporaneous effect on growth revenues and so they can be considered as the driver process of firm growth. Moreover, it affects positively all the other variables and if we look the overall picture of causal structure inferred from the previous SVAR model, it is evident how the leading process of firm growth is actually the revenues. This is well explained in the causal graph of Fig. 1.1.

In conclusion, the analysis of this chapter empirically motivates the conventional wisdom that the revenues constitute the driver process of firm growth. Therefore, in order to construct a stochastic valuation model relying on the revenues dynamics, in Chapter 3 we will express all the other accounting variables which define the UFCF, see

[13] We test also the non-Gaussianity for structural shocks and we obtain analogous results of those of the reduced ones.

Table 1.3 Coefficients of the instantaneous effects matrix B and standard errors, in parenthesis, calculated from 100 bootstrap performed on residuals. The coefficients in bold are significantly different from zero using a t-test at significance level 1%.

	Rev.gr	EBITDA.gr	DA.gr	CAPEX.gr
Rev.gr	0	-0.044	0.4846	-2.7421
	-	(0.1551)	(0.4885)	(2.519)
EBITDA.gr	**0.6511**	0	-0.0231	-0.0886
	(0.0322)	-	(0.0094)	(0.0403)
DA.gr	**0.4567**	**0.2467**	0	0.0236
	(0.0841)	(0.0719)	-	(0.0777)
CAPEX.gr	**0.525**	**0.1616**	**0.1281**	0
	(0.0173)	(0.0203)	(0.0052)	-

Table 1.4 Coefficients of the lagged effects matrix A_1^* and standard errors, in parenthesis, calculated from 100 bootstrap performed on residuals. The coefficients in bold are significantly different from zero using a t-test at significance level 1%.

	Rev.gr	EBITDA.gr	DA.gr	CAPEX.gr
Rev.gr	**0.1135**	0.0746	0.0064	-0.1491
	(0.0342)	(0.0977)	(0.008)	(0.1286)
EBITDA.gr	**-0.0184**	-0.003	-0.0009	0.004
	(0.0066)	(0.0035)	(0.0012)	(0.0047)
DA.gr	-0.020	**0.0278**	**-0.0184**	-0.0223
	(0.0102)	(0.0081)	(0.0038)	(0.0112)
CAPEX.gr	**0.0182**	-0.0013	0.003	**-0.0144**
	(0.0037)	(0.0028)	(0.0012)	(0.0033)

Eq. (1.3), as margin on sales. However, in the next chapter we focus on the above ICA methodology employed for the structural analysis, where we present a general setting and related extensions of this identification approach.

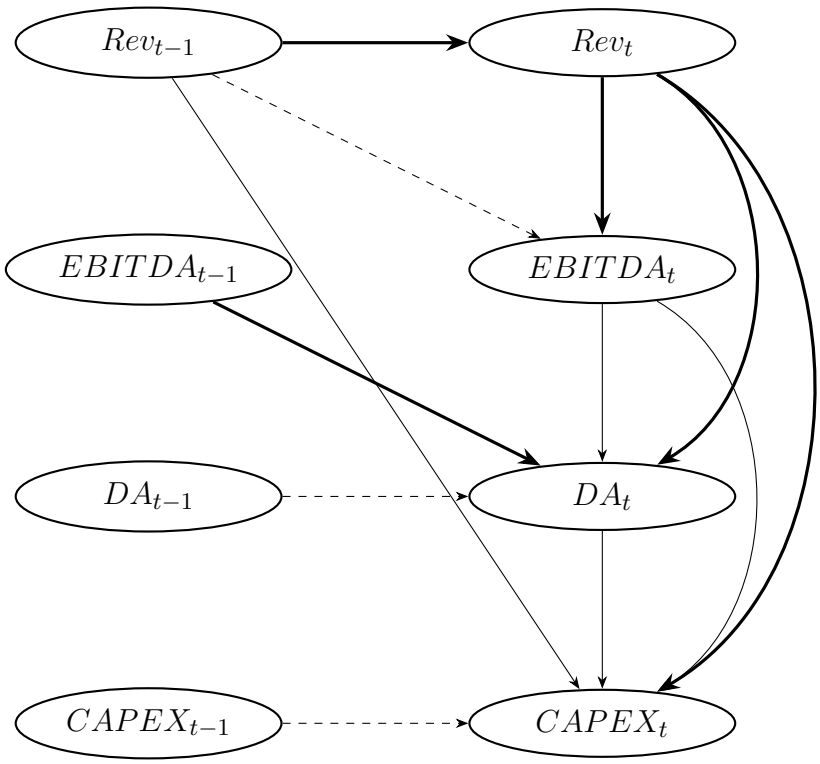


Fig. 1.1 Causal graph inferred from the estimated SVAR model. Only significant values of B and Γ_1 are considered. Solid (dashed) arrows indicate positive (negative) effects. Thick (thin) lines correspond to strong (weak) effects. An effect is classified as strong when its absolute value is greater to the median of all significative effects (in absolute value).

CHAPTER 2

Identification of Singular and Noisy Structural VAR Models

We now introduce extensions of the ICA approach for the SVAR identification problem. When the number of observed variables is larger than the number of structural shocks the associated structural VAR system is said to be singular. We propose an identification method for singular structural VAR models contaminated by noise that combines a collapsing procedure with the independent component analysis.

The need to extract information from a large number of macroeconomic series, e.g. to deal with the invertibility problem, (see Forni et al. (2009) and Stock and Watson (2016)), leads to singular SVAR systems, i.e., the variance-covariance matrix of the reduced shocks is singular. This could emerge also in the context of sign restrictions, as pointed by Granziera et al. (2018). Moreover, the singular case naturally arises in the context of Dynamical Factor Models (henceforth DFM), where the structural analysis can be conducted directly on the common factors, which are driven generally by a fewer structural shocks. In particular, the DFM residuals achieve the fundamentalness property, which is not guaranteed to be satisfied by the structural shocks obtained by a SVAR model, see Forni et al. (2009), but thank to the results of Anderson and Deistler (2008b), we show that when a SVAR model is singular the structural shocks turn out to be fundamental and can be identified from the reduced ones.

However, it seems reasonable to assume that the observed signal, the residuals of the reduced form model, in addition to be a mixing of the unobserved signal given by the structural shocks, might be contaminated by different types of noises: measurement errors on the the macro-variables, estimation errors on the parameter of the reduced form model and model misspecification of the VAR. For instance, Forni et al. (2017) examine a model of imperfect information where the agents observe a noisy signal and they propose an identification method based on dynamic rotations where agents learn by future information the composition of the structural shocks. In contrast to our approach, they interpret the noise as a source of business cycle fluctuations, while in our framework the noise shock is interpreted as a measurement error which affects the macroeconomic variables as in the standard DFM framework, see Forni et al. (2000) and Forni et al. (2009). Hence, we assume that some noise contaminating the system remains, thus leading to a problem of noisy ICA. We discuss under which assumptions the combined

procedure remains feasible when the number of time observations gets large.

We first introduce the extensions of standard ICA problem in the SVAR context, then we show how to combine the collapsing procedure with the ICA technique discussing under which assumptions the proposed procedure is consistent and feasible. In contrast to the previous exercise on the microeconomic study for firm growth, we analyze the proposed scheme with an empirical macroeconomic application on U.S. data to study the effect of monetary policy[1]. The new identification method allows to employ the entire set of information contained in all the observed macro-variables available to consistently identifying the lower dimensional system of the structural shocks driving the U.S. economy. Finally, in Appendix A.2 we report the sensitivity analysis simulation results of the method and we also examine its finite sample properties with Monte Carlo simulations.

Let us recall some notation. If $\boldsymbol{y}_t \in \mathbb{R}^N$ is the vector of the observable time series of interest a VAR(p) model is given by the following equation:

$$\boldsymbol{y}_t = \boldsymbol{v} + A_1 \boldsymbol{y}_{t-1} + \cdots + A_p \boldsymbol{y}_{t-p} + \boldsymbol{u}_t, \tag{2.1}$$

where, $\boldsymbol{v} \in \mathbb{R}^N$ is the intercept, $A_1, A_2, \ldots, A_p \in \mathbb{R}^{N\times N}$ are parameter matrices and $\boldsymbol{u}_t \in \mathbb{R}^N$ are the residuals which are serially uncorrelated with mean zero and covariance matrix $\Sigma_{\boldsymbol{u}}$. Then, let $\boldsymbol{\alpha}_t \in \mathbb{R}^N$ be the vector of structural shocks, which are temporally uncorrelated, strictly stationary, with zero mean, uncorrelated $\mathrm{Cov}(\boldsymbol{\alpha}_t) = \Sigma_{\boldsymbol{\alpha}} = I_N$ and such that $\boldsymbol{u}_t = Z\boldsymbol{\alpha}_t$, where $Z \in \mathbb{R}^{N\times N}$ is nonsingular and it is satisfies $\Sigma_{\boldsymbol{u}} = ZZ'$. So, since stationary time series are considered, i.e., $\det A(z) := \det(I_N - A_1 z - \cdots - A_p z^p) \neq 0,\ |z| \leq 1,\ z \in \mathbb{C}$, we can obtain the infinite moving average representation of the VAR,

$$\boldsymbol{y}_t = \boldsymbol{\mu} + \sum_{k=0}^{\infty} \Psi_k Z \boldsymbol{\alpha}_{t-k}, \quad \Psi_0 = I_N,$$

where $\boldsymbol{\mu}$ is the expectation of $\boldsymbol{y}_t$ and $A(z)^{-1} = \Psi(z) = \sum_{k=0}^{\infty} \Psi_k z^k$. Therefore, the Impulse Response Function (IRF) of $y_{i,t}$ to a unitary shock on $\alpha_{j,t}$ is given by:

$$IRF_{i,j}(k) = \Psi_k(i,:) \cdot Z(:,j).$$

Consistently with Section 1.2.3, we refer to the structural, $\boldsymbol{\alpha}_t$, and reduced shocks, $\boldsymbol{u}_t$, as sources and sensors, respectively.

2.1 Extensions of the ICA approach to SVAR models

So far, the ICA technique has been employed in the SVAR identification problem when the number of the structural shocks, R, is equal to the number of observed variables, N. We propose two types of generalizations: the ICA problem is first extended to the singular

[1]This is the standard application for SVAR analysis since Sims (1980).

case and then the noisy formulation is introduced. In the ICA literature the singular case is also named *overdetermined* or *undercomplete*, however we use also the denomination of singular ICA when the associated VAR model is singular.

2.1.1 Singular ICA

Solving a singular (overdetermined) ICA problem amounts in finding the mixing matrix Z and the source $\boldsymbol{\alpha}$ given the observations $\boldsymbol{u}$, such that:

$$\boldsymbol{u}_t = Z\boldsymbol{\alpha}_t,\ \boldsymbol{u}_t \in \mathbb{R}^N,\ \boldsymbol{\alpha}_t \in \mathbb{R}^R,\ Z \in \mathbb{R}^{N\times R} \text{ where } N > R.$$

In other words, we have more observations than sources. Temporarily, we may assume that the number of sources is known a priori[2], nonetheless it can be estimated by using the consistent estimator for select the number of factors following Bai and Ng (2002), Hallin and Liška (2007), Alessi et al. (2010) and Onatski (2010). In the general ICA context of blind source separation the overdetermined problem has been investigated in Hyvärinen et al. (2001), Joho et al. (2000). This problem naturally arises when the associated SVAR is singular, i.e., the number of reduced shocks is greater than the number of structural shocks, so that the variance-covariance matrix $\boldsymbol{\Sigma_u}$ becomes singular. For instance, when in a simple VAR framework different macroeconomic variables are employed as measures of the same economic variable, so that they are concerned to the same structural shock, the associated SVAR model becomes singular. In Section 2.3 we study an empirical application where we use two different proxies for the economic activity [3].

2.1.2 Noisy ICA

The presence of parameter and model uncertainty in the VAR setting, imperfect information and measurement errors in the macroeconomic variables, suggests to assume that some form of noise remains in the reduced shocks, leading to a noisy ICA problem. In other words, the VAR can be seen as an imperfect filter on the observed variables $\boldsymbol{y}_t$, so that it seems realistic to assume that some form of noise remains in the sensors. The noisy ICA problem can be formulated in two ways. We can either assume that the noise is generated directly from $\boldsymbol{\alpha}_t$ (*source noise* formulation) or that it is added to the mixture $Z\boldsymbol{\alpha}_t$ (*sensor noise* formulation). The noisy ICA formulation was studied by Moulines et al. (1997), Attias (1999), Hyvärinen (1999), Ikeda (2000), Bonhomme and Robin (2009) as an extension of a standard blind source separation problem.

[2] In the following we discuss how to derive this number.

[3] Where we also perform principal component analysis to verify that the associated VAR problem is effectively singular.

2.1.2.1 Source noise

If each source generates a noise term, then the noise $\boldsymbol{e}_t^{\boldsymbol{\alpha}}$, is added directly to the source: $\boldsymbol{\beta}_t = \boldsymbol{\alpha}_t + \boldsymbol{e}_t^{\boldsymbol{\alpha}}$. The noise $\boldsymbol{e}_t^{\boldsymbol{\alpha}}$ is assumed Gaussian and independent from $\boldsymbol{\alpha}_t$, so that the sources $\boldsymbol{\beta}_t$ are independent and such that $\mathbb{E}[\boldsymbol{e}_t^{\boldsymbol{\alpha}}\boldsymbol{e}_t^{\boldsymbol{\alpha}'}] = \sigma^2 I_R$, consistently with Assumptions 1.1. However:

$$\boldsymbol{u}_t = Z\boldsymbol{\beta}_t = Z\boldsymbol{\alpha}_t + Z\boldsymbol{e}_t^{\boldsymbol{\alpha}} = Z\boldsymbol{\alpha}_t + \boldsymbol{e}_t$$

where $\boldsymbol{u}_t \in \mathbb{R}^N$, $\boldsymbol{\alpha}_t, \boldsymbol{e}_t^{\boldsymbol{\alpha}} \in \mathbb{R}^R$, $Z \in \mathbb{R}^{N\times R}$ and $\boldsymbol{e}_t := Z\boldsymbol{e}_t^{\boldsymbol{\alpha}} \in \mathbb{R}^N$. So, we can identify Z by considering the overdetermined ICA problem $\boldsymbol{u}_t = Z\boldsymbol{\beta}_t$ as for the noiseless case, since also the components of $\boldsymbol{\beta}_t$ are non-Gaussian and independent. In the following, we discuss the noiseless case when we describe the Collapsing-ICA procedure and we show how it is related to the particular case of source noise.

2.1.2.2 Sensor noise

Under the assumption that the noise term is generated by the sensors, the noisy ICA problem can be formulated as

$$\boldsymbol{u}_t = Z\boldsymbol{\alpha}_t + \boldsymbol{e}_t, \tag{2.2}$$

where $\boldsymbol{u}_t \in R^N$ is the sensor signal, $\boldsymbol{\alpha}_t \in \mathbb{R}^R$ is the source signal, $Z \in \mathbb{R}^{N\times R}$ is the mixing matrix and $\boldsymbol{e}_t \in \mathbb{R}^N$ is an error term, the noise. Essentially, the noise term $\boldsymbol{e}_t$ is added to the mixture of the sources $Z\boldsymbol{\alpha}_t$. Standard assumptions, see e.g. Hyvärinen et al. (2001), are: i) the noise is independent from the sources and ii) the noise is mutually independent and Gaussian with an unknown variance-covariance matrix H. In Section 2.2.3.2 and Appendix A.1 we discuss how we can relax the assumption of mutually independent noise and how we can deal with a generic unknown covariance matrix H.

Remark 2.1. As highlighted by Hyvärinen et al. (2001), under suitable assumptions this two noise formulations are equivalent. Indeed, if $\text{Cov}(\boldsymbol{e}_t) = ZH_1Z'$ where $H_1 \in \mathbb{R}^{R\times R}$, let $\boldsymbol{e}_t = Z\boldsymbol{e}_t^{\boldsymbol{\alpha}}$ where $\boldsymbol{e}_t^{\boldsymbol{\alpha}} \in \mathbb{R}^R$, $\text{Cov}(\boldsymbol{e}_t^{\boldsymbol{\alpha}}) = H_1$, then

$$\boldsymbol{u}_t = Z\boldsymbol{\alpha}_t + Z\boldsymbol{e}_t^{\boldsymbol{\alpha}} = Z(\boldsymbol{\alpha}_t + \boldsymbol{e}_t^{\boldsymbol{\alpha}}) = Z\boldsymbol{\beta}_t, \quad \boldsymbol{\beta}_t = \boldsymbol{\alpha}_t + \boldsymbol{e}_t^{\boldsymbol{\alpha}},$$

where $\boldsymbol{\beta}_t$ denotes the contaminated sources.

The source noise formulation implicitly assumes that a structural shock is composed by a "true" shock and a noise shock, which is very much in line with what Forni et al. (2017) assume. Indeed, they consider that the fundamental shock is the sum of a "long-run" shock and a noise shock, where this noise shock is interpreted as a source of business cycle fluctuations. Although the source noise formulation can be easily solved by the same techniques for the noiseless case, it is more reasonable to consider the case in which the

noise is added to the mixture of the sources instead of assume that each structural shocks generate a noise term. Indeed, the structural shocks can be viewed as "pure" signals while their measurement is affected by a noise. Thus, in the following, we consider the noisy ICA with the second formulation, where the noise should be interpreted as a measurement error which is present in the macroeconomic variables in addition to the model and parameter uncertainty of the VAR setting. In other words, we study the singular VAR model (2.1) where the reduced shocks are described by Eq. (2.2).

The resulting model is related with the Factor Augmented VAR (FAVAR) model considered in Bernanke et al. (2005) and Factor-Structural Vector Autoregression (FSVAR) model described in Stock and Watson (2005). However, FAVAR method can be interpreted as an extension of the DFM approach where, as observed by Stock and Watson (2016), restrictions are imposed on the DFM, i.e., that one or more of the factors is treated as observed without measurement error, while our approach is an extension of the VAR approach without requiring any type of restrictions on the reduced model. On the other hand, the FSVAR model imposes a factor structure to the reduced shocks. It is estimated by Gaussian maximum likehood, but, in contrast to Stock and Watson (2005), we assume that the structural shocks are non-Gaussian (and independent) so that identification of the mixing matrix is achieved without any type of economic assumptions/restrictions[4] using ICA. Moreover, the assumption of non-Gaussian shocks seems to be much in line with the statistical proprieties of VAR residuals, as we observe from our empirical results and as also supported by the evidence of Moneta et al. (2013), Gouriéroux et al. (2017), Lanne et al. (2017).

In the next section we explore how to combine the collapsing procedure of Jungbacker and Koopman (2008) and the PML approach to estimate Z. We will discuss the approach feasibility, the identification and fundamentalness issues together with consistency when T, the number of time observations, goes to infinity. However, the presented procedure remains valid for any other consistent ICA approach which can be used alternatively to the PML estimator.

2.2 The Collapsing-ICA procedure

In this section we illustrate how to exploit the collapsing technique of Jungbacker and Koopman (2008) for the general singular noisy ICA problem. We assume the same model of Eq. (2.2) for the residuals of a VAR model: $\boldsymbol{u}_t = Z\boldsymbol{\alpha}_t + \boldsymbol{e}_t$, $\boldsymbol{e}_t \sim N(0, H)$, where for each t, $\boldsymbol{u}_t \in \mathbb{R}^N$ are the reduced shocks, $\boldsymbol{\alpha}_t \in \mathbb{R}^R$ are the structural shocks, $Z \in \mathbb{R}^{N\times R}$ is the mixing matrix and $\boldsymbol{e}_t \in \mathbb{R}^N$ is the noise which affects the mixture of structural shocks. We want to reduce the dimension of the problem in such a way that also the noise term is reduced. The collapsing procedure of Jungbacker and Koopman (2008) replaces

[4] Stock and Watson (2005) identify the factors by assuming that the second factor has no impact effect on one of the considered variables.

$\boldsymbol{u}_t \in \mathbb{R}^N$ with the so called "virtual" or collapsed observations $\boldsymbol{u}_t^1 \in \mathbb{R}^R$ by employing a projection matrix $\Pi_1 \in \mathbb{R}^{R\times N}$, which is defined by $\Pi_1 = (Z'H^{-1}Z)^{-1}Z'H^{-1} \in \mathbb{R}^{R\times N}$. Then, we also consider $\Pi_2 = B(I_N - Z\cdot\Pi_1) \in \mathbb{R}^{(N-R)\times N}$, where B is such that Π_2 has full rank of $N-R$, so that, $\Pi_1 Z = I_R$ and $\Pi_2 Z = 0$. So, if $\boldsymbol{u}_t^1 = \Pi_1\boldsymbol{u}_t$, $\boldsymbol{u}_t^2 = \Pi_2\boldsymbol{u}_t$, $\boldsymbol{e}_t^1 = \Pi_1\boldsymbol{e}_t$ and $\boldsymbol{e}_t^2 = \Pi_2\boldsymbol{e}_t$,

$$\Pi^*\boldsymbol{u}_t = \begin{pmatrix}\boldsymbol{u}_t^1\\ \boldsymbol{u}_t^2\end{pmatrix} = \begin{pmatrix}\boldsymbol{\alpha}_t\\ 0\end{pmatrix} + \begin{pmatrix}\boldsymbol{e}_t^1\\ \boldsymbol{e}_t^2\end{pmatrix}, \text{ where } \Pi^* = \begin{pmatrix}\Pi_1\\ \Pi_2\end{pmatrix} \in \mathbb{R}^{N\times N},$$

$\boldsymbol{e}_t^1$ and $\boldsymbol{e}_t^2$ are independent, since they are Gaussian and $\mathrm{Cov}(\boldsymbol{e}_t^1, \boldsymbol{e}_t^2) = 0$, and their variance-covariance matrices are given by $H_1 = \Pi_1 H \Pi_1' = (Z'H^{-1}Z)^{-1}$ and $H_2 = \Pi_2 H \Pi_2'$, respectively. Hence, all the information is collapsed in $\boldsymbol{u}_t^1$, to which we can apply the ICA step directly[5] .

The general Collapsing-ICA (C-ICA) procedure is summarized in the following scheme. Let $Z = U\Sigma C'$ be a singular value decomposition (SVD)[6] of Z, where $U \in \mathbb{R}^{N\times N}$ and $C' \in \mathbb{R}^{R\times R}$ are orthogonal matrices[7] and $\Sigma \in \mathbb{R}^{N\times R}$ is the matrix with the singular values of Z. We may consider the product $U\Sigma$ as an initial estimate for Z and once obtained U, Σ and H, the collapsing procedure can be applied on $\boldsymbol{u}_t$. Then the ICA is employed on the collapsed equation in order to estimate C', see Algorithm 1. The estimate of U and Σ can be obtained, for instance, by performing a PCA on $\boldsymbol{u}_t$. However, this initialization step depends on the assumptions of noise and it is discussed together with the estimate of H in the next sections.

Algorithm 1 The general C-ICA procedure. Let $Z = U\Sigma C'$ be a singular value decomposition of the mixing matrix.

Initialization: Get an estimate† of the system matrices U, Σ and H.

1: *Collapsing-Step*: Use the collapsing procedure and obtain the collapsed observations, $\boldsymbol{u}_t^1 = \Pi_1\boldsymbol{u}_t$.
2: *ICA-Step*: Perform ICA on $\boldsymbol{u}_t^1$ and estimate C'.

†The estimation of the system matrices is discussed in the following subsections together with the required assumptions for consistency.

We remark that $U\Sigma$ can not be used for the computation of the IRFs, since the

[5]We observe that the two step procedure of Joho et al. (2000) is a particular case of the collapsing procedure.

[6]Even if, the standard notation of the SVD of a matrix is $Z = U\Sigma V'$, coherently with the SVAR literature we denote V as C.

[7]For notational convenience and simplicity, we will denote the orthogonal matrix of the standard PML-ICA approach with the same notation used for the C-ICA. Thus, the decomposition employed by the PML-ICA is rewritten as $Z = SC'$, where S is the Cholesky decomposition of $\Sigma_{\boldsymbol{u}}$ and C' orthogonal. Since both C' matrices of C-ICA and PML-ICA play the same role in describing the mixing process, we prefer to unify the notations consistently with the presented SVD decomposition.

structural shocks are not separated yet. Only after the ICA step we can analyze the responses of each variable to the respective shock. In other words, the initial estimate of Z is up to an orthogonal matrix, which is represented by the matrix C' when we consider the singular value decomposition of Z.

2.2.1 Identification results

We now present preliminary results on the identification with the C-ICA procedure and then we show the consistency of our procedure for the noiseless case and noisy case. The following lemmas are particular cases of Autonne's uniqueness Theorem, see Theorem 2.5.6 of Horn and Johnson (1990) for a complete discussion. Let us consider a SVD of $Z \in \mathbb{R}^{N\times R}$, $Z = U\Sigma C'$. The matrix of the singular values is uniquely identify by Z, while the orthogonal matrix U and C are not. However, we can find a relation between different pairs of U and C. We first analyze the particular case of $N = R$.

Lemma 2.2. *Let $Z \in \mathbb{R}^{N\times N}$ and let us consider a singular value decomposition of Z, $Z = U_1\Sigma C_1'$ where U_1, C_1 are $N \times N$ orthogonal matrices and Σ is the diagonal matrix which contains the singular values of Z. If the singular values of Z are distinct then, given two singular value decompositions*

$$Z = U_1\Sigma C_1' = U_2\Sigma C_2',$$

the matrices U_1 and C_1 are identified up to a multiplication by the same orthogonal diagonal matrix B, i.e.,

$$U_2 = U_1B,\ C_2 = C_1B.$$

Proof of Lemma 2.2. Starting from, $Z'Z = C_1\Sigma'\Sigma C_1' = C_2\Sigma'\Sigma C_2'$ if we multiply both sides for C_1' and C_2, left and right respectively, we obtain $\Sigma'\Sigma \cdot C_1'C_2 = C_1'C_2 \cdot \Sigma'\Sigma$. Then, if we denote by $A = \Sigma'\Sigma$ and $B = C_1'C_2$ we observe that A and B are simultaneously diagonalizable, since they commute and they are both diagonalizable. Furthermore, A is a diagonal matrix with distinct elements which implies that also B is diagonal. More precisely, B is an orthogonal diagonal matrix and we have that $C_2 = C_1B$. Using this relation on the original SVD, we get $Z = U_1\Sigma C_1' = U_2\Sigma(C_1B)' = U_2\Sigma B'C_1' = U_2B'\Sigma C_1'$ and since Σ is squared we have also that $U_2 = U_1B$. □

Remark 2.3. The diagonal elements of B can only be ± 1, since B is an orthogonal diagonal matrix. Then, the identification of U and C is up to change of signs.

Let us now consider the general case where $N > R$.

Proposition 2.4. *Let $Z \in \mathbb{R}^{N\times R}$, which has full rank R, and let us consider a singular value decomposition of Z, $Z = U_1\Sigma C_1'$ where $U_1 \in \mathbb{R}^{N\times N}$, $C_1 \in \mathbb{R}^{R\times R}$ are orthogonal matrices and $\Sigma = \begin{bmatrix}\Sigma_+ \\ 0\end{bmatrix} \in \mathbb{R}^{N\times R}$ where Σ_+ is the $R \times R$ diagonal matrix which contains*

the singular values of Z. If the singular values of Z are distinct then, given two singular value decompositions

$$Z = U_1 \Sigma C_1' = U_2 \Sigma C_2',$$

the first R columns of U_1 and the matrix C_1 are identified up to a multiplication by the same orthogonal diagonal matrix B, i.e.,

$$U_{2_R} = U_{1_R} B, \quad C_2 = C_1 B,$$

where U_{1_R} and U_{2_R} are the first R columns of U_1 and U_2, respectively.

Proof of Proposition 2.4. Let us start again by considering, $Z'Z = C_1 \Sigma' \Sigma C_1' = C_2 \Sigma' \Sigma C_2'$. We observe that $A = \Sigma' \Sigma \in \mathbb{R}^{R \times R}$ is a diagonal matrix and we can use the same argument as for the proof of Lemma 2.2 to obtain that $C_2 = C_1 B$, where $B \in \mathbb{R}^{R \times R}$ is an orthogonal diagonal matrix. Then,

$$\begin{aligned} Z = U_1 \Sigma C_1' &= U_2 \Sigma (C_1 B)' \\ &= U_2 \Sigma B' C_1' \\ &= U_2 \begin{bmatrix} B' \\ 0 \end{bmatrix} \Sigma_+ C_1', \end{aligned} \tag{2.3}$$

where we use that $\Sigma B' = \begin{bmatrix} \Sigma_+ \\ 0 \end{bmatrix} B' = \begin{bmatrix} \Sigma_+ B' \\ 0 \end{bmatrix} = \begin{bmatrix} B' \Sigma_+ \\ 0 \end{bmatrix} = \begin{bmatrix} B' \\ 0 \end{bmatrix} \Sigma_+$, since B' and Σ_+ are diagonal matrices. Then, if we multiply both right sides of Eq. (2.3) by $C_1 \Sigma_+^{-1}$, we obtain

$$U_1 \begin{bmatrix} I_R \\ 0 \end{bmatrix} = U_2 \begin{bmatrix} B' \\ 0 \end{bmatrix},$$

and if $U_{i_R} \in \mathbb{R}^{N \times R}$ denotes the first R columns of U_i for $i = 1, 2$, we finally obtain that $U_{2_R} = U_{1_R} B$. □

We conclude by considering the particular case when $\Sigma_+ = \sigma I_R$.

Lemma 2.5. *Let $Z \in \mathbb{R}^{N \times R}$, which has full rank R, and given a singular value decomposition of Z, $Z = U \Sigma C'$ where $U \in \mathbb{R}^{N \times N}$, $C' \in \mathbb{R}^{R \times R}$ are orthogonal matrices and $\Sigma = \begin{bmatrix} \sigma I_R \\ 0 \end{bmatrix} \in \mathbb{R}^{N \times R}$. Given the first R columns of U, denoted by U_R, then C is uniquely identified.*

Proof of Lemma 2.5. If there exists C_2, an orthogonal matrix such that, $Z = \sigma U_R C_1' = \sigma U_R C_2'$, then $U_R = U_R C_2' C_1$, but $U_R \in \mathbb{R}^{N \times R}$ and it satisfies $U_R' U_R = I_R$, so $I_R = C_2' C_1$ which implies $C_2 = C_1$. □

In the following, we discuss under which assumptions the C-ICA approach is consistent.

2.2.2 The noiseless case

We first describe the C-ICA procedure in the particular case when the VAR is only singular, i.e., when in Eq. (2.2) there is no noise. Let $Z = U\Sigma C'$ be a SVD of Z. We assume that the matrix Z is of full rank, i.e., its singular values are different from zero. Therefore, when $T \to \infty$, the variance-covariance of $\boldsymbol{u}_t$, $\Sigma_{\boldsymbol{u}} = U\Sigma\Sigma' U'$, is singular and the number of structural shocks R can be obtained by computing its rank. The initialization step of C-ICA procedure consists in the computation of the first R columns of U, U_R, and the singular values $\Sigma_+ = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_R)$ by using a spectral decomposition on $\Sigma_{\boldsymbol{u}}$. Then, the projection matrix, of the collapsing step, is defined as $\Pi_1 = \Sigma_+^{-1} U_R' \in \mathbb{R}^{R\times N}$, and it is straightforward that $\boldsymbol{u}_t^1 = \Pi_1 \boldsymbol{u}_t = C'\boldsymbol{\alpha}_t$. Thus, any consistent ICA technique can be applied to $\boldsymbol{u}_t^1$ to obtain a consistent estimator of C' and by Proposition 2.4, and Lemma 2.5, we known that U_R and C' are identified up to an orthogonal diagonal matrix. We summarize the previous results in the following Proposition, where the PML estimator of Gouriéroux et al. (2017) is used in the ICA-step. Since we employ the PML estimator, the technical Assumptions A.2-A.5 of Gouriéroux et al. (2017) are necessary in order to obtain consistency of their estimator. These assumptions ensure that the pseudo-density functions of the estimator are sufficiently regular and they are distinct and asymmetric as well as the true distributions of $\boldsymbol{\alpha}_t$. We refer to Section 2.3 of Gouriéroux et al. (2017) for a complete discussion.

Proposition 2.6. *Given $\boldsymbol{u}_t = Z\boldsymbol{\alpha}_t$, where $Z \in \mathbb{R}^{N\times R}$ has full rank $R, \boldsymbol{u}_t \in \mathbb{R}^N, \boldsymbol{\alpha} \in \mathbb{R}^R$. If the singular values of Z are distinct, or they are all equal, and $\boldsymbol{\alpha}_t$ satisfies Assumptions 1.1, then, under Assumptions A.2-A.5 of Gouriéroux et al. (2017), the C-ICA procedure, where the PML estimator of Gouriéroux et al. (2017) is employed in the ICA-step, is consistent.*

Proof of Proposition 2.6. Let us consider the variance-covariance matrix of $\boldsymbol{u}_t$, $\mathbb{E}[\boldsymbol{u}_t \boldsymbol{u}_t'] = \Sigma_{\boldsymbol{u}}$. By the SVD of $Z = U\Sigma C'$ we obtain when $T \to \infty$ that $\Sigma_{\boldsymbol{u}} = U\Sigma C'\Sigma_{\boldsymbol{\alpha}} C\Sigma' U' = U\Sigma\Sigma' U'$, since $\Sigma_{\boldsymbol{\alpha}} = I_R$, then

$$\Sigma_{\boldsymbol{u}} = U \begin{bmatrix} \Sigma_+^2 & 0 \\ 0 & 0 \end{bmatrix} U'.$$

If we denote with $U_R \in \mathbb{R}^{N\times R}$ the first R columns of U and we observe that ${U_R}' U_R = I_R$. Then, using a spectral decomposition on $\Sigma_{\boldsymbol{u}}$ we can compute U_R and Σ_+. So, the collapsing procedure, where $\Pi_1 = \Sigma_+^{-1} {U_R}' \in \mathbb{R}^{R\times N}$, leads to $\boldsymbol{u}_t^1 = \Pi_1 \boldsymbol{u}_t = \Sigma_+^{-1} U_R' U\Sigma C'\boldsymbol{\alpha}_t = \Sigma_+^{-1} \left[I_R \mid 0 \right] \Sigma C'\alpha_t = C'\boldsymbol{\alpha}_t$. Then, the PML estimator employed on $\boldsymbol{u}_t^1$ results to be a consistent estimator of C', where the consistency follows from the same argument of

Gouriéroux et al. (2017). The identifiability of U_R and C' is guaranteed by Proposition 2.4 and Lemma 2.5. □

The assumptions on the singular values are necessary for the identification, see Proposition 2.4 and Lemma 2.5. The procedure is consistent in the sense that C' is consistently estimated up to permutation and change of signs, which means that the structural analysis can be conducted on $\boldsymbol{\alpha}_t$. In addition, the estimator of C' has the asymptotic properties described in Gouriéroux et al. (2017). In the following we refer to Σ_+ as the amplification matrix and C' as the collapsed mixing matrix.

2.2.2.1 The Fundamentalness of Structural Shocks in Singular VAR Models

In the general setting the structural shocks of a non-singular VAR model are non fundamental. Gouriéroux et al. (2019) have shown that this problem is strictly related to the Gaussian assumption of the structural shocks and that if Assumptions 1.1 holds this difficulty is overcome. However, if the VAR is singular the fundamentalness property is guaranteed without any further assumption thank to the results of Anderson and Deistler (2008a,b).

We recall the definition of fundamental shock. The fundamentalness is a generic property in the sense of Forni et al. (2009), which we refer for further details.

Definition 2.7. The vector of shocks $\boldsymbol{\alpha}_t$ is *fundamental* for $\boldsymbol{y}_t$ if

- there exists a moving average representation $\boldsymbol{y}_t = B(L)\boldsymbol{\alpha}_t$ and
- $\boldsymbol{\alpha}_t \in \overline{\text{Span}}(\boldsymbol{y}_t, \boldsymbol{y}_{t-1}, \ldots)$.

Proposition 2.8. *Let us consider the reduced form of a VAR model $A(L)\boldsymbol{y}_t = \boldsymbol{u}_t$, $\boldsymbol{y}_t \in \mathbb{R}^N, \boldsymbol{u}_t \in \mathbb{R}^N$, where L is the lag operator. Let $\boldsymbol{y}_t$ satisfies Assumptions 1-5 of Anderson and Deistler (2008a). If we assume that the (S)VAR is singular, i.e., $\boldsymbol{u}_t = Z\boldsymbol{\alpha}_t$, $\boldsymbol{\alpha}_t \in \mathbb{R}^R$, $Z \in \mathbb{R}^{N\times R}$, $N > R$, where Z has full rank R and $\boldsymbol{\alpha}_t$ is the vector of uncorrelated structural shocks, then $\boldsymbol{\alpha}_t$ is fundamental. Moreover, $\boldsymbol{\alpha}_t$ can be identified from the observed reduced shocks $\boldsymbol{u}_t$.*

Proof of Proposition 2.8. From the observed residuals $\boldsymbol{u}_t$ we recover the associated uncorrelated fundamental shocks $\boldsymbol{\beta}_t \in \mathbb{R}^R$, which are not generically the structural ones, by using the relation $\boldsymbol{u}_t = W\boldsymbol{\beta}_t$, for a certain full-rank matrix $W \in \mathbb{R}^{N\times R}$. $\boldsymbol{\beta}_t$ is fundamental, since $\boldsymbol{u}_t$ is by construction fundamental. Then, we may recover the moving average representation

$$\boldsymbol{y}_t = B(L)W\boldsymbol{\beta}_t. \tag{2.4}$$

On the other hand, if we consider the structural representation of the reduced VAR model

$$\boldsymbol{y}_t = C(L)Z\boldsymbol{\alpha}_t, \tag{2.5}$$

we observe that since the VAR is singular, from the results of Anderson and Deistler (2008b) and Theorem 3 of Anderson and Deistler (2008a) $\boldsymbol{\alpha}_t$ is generically fundamental. In particular, since $\boldsymbol{\alpha}_t$ and $\boldsymbol{\beta}_t$ are both fundamental shocks, from Eq. (2.4) and (2.5) we obtain that $\boldsymbol{\beta}_t$ and $\boldsymbol{\alpha}_t$ are the same up to an orthogonal matrix, (see Proposition 2 of Forni et al. (2009)), which means that we can identity $\boldsymbol{\alpha}_t$ from the observed $\boldsymbol{u}_t$. □

Assumptions 1-5 of Anderson and Deistler (2008a) ensure the equivalence of having a generic zeroless MA representation of $\boldsymbol{y}_t$ to the singular case, see Anderson and Deistler (2008a) for technical details.

In particular, the (fundamental) shocks that we can recover from the observed $\boldsymbol{u}_t$ shocks differ from the structural ones $\boldsymbol{\alpha}_t$ only for an orthogonal transformation Q. This means that under the assumptions of Proposition 2.6 we can identify from the reduced observed shocks $\boldsymbol{u}_t$ the structural ones which are fundamental, i.e., Q is a permutation and change of signs.

2.2.3 The noisy case

We now deal with the noisy case. When the noise is such that $H = \Sigma_{\boldsymbol{e}} = \mathbb{E}[\boldsymbol{e}_t \boldsymbol{e}_t'] = \sigma_{\boldsymbol{e}}^2 I_N$ we refer to it as *homoskedastic* noise and otherwise as *heteroskedastic*.

2.2.3.1 The homoskedastic noise case

We first consider the homoskedastic noise case. We note that this assumption is also adopted by Stock and Watson (2005). We first assume that the amplification matrix is homogeneous and then we discussed the consequences of relaxing the adopted assumptions.

Proposition 2.9. *Given*

$$\boldsymbol{u}_t = Z\boldsymbol{\alpha}_t + \boldsymbol{e}_t, \text{ where } Z \in \mathbb{R}^{N\times R} \text{ and } rank(Z) = R, \quad \boldsymbol{u}_t, \boldsymbol{e}_t \in \mathbb{R}^N, \boldsymbol{\alpha} \in \mathbb{R}^R.$$

Let $Z = U\Sigma C'$ be a singular value decomposition of Z, where $U \in \mathbb{R}^{N\times N}, C' \in \mathbb{R}^{R\times R}$ are orthogonal matrices and $\Sigma = \begin{bmatrix} \Sigma_+ \\ 0 \end{bmatrix}$. Let us suppose:

- $\boldsymbol{e}_t \sim N(0, H)$, $H = \sigma_{\boldsymbol{e}}^2 I_N$ *where* $\boldsymbol{e}_t$ *is independent from* $\boldsymbol{\alpha}_t$,
- $\Sigma_+ = \sigma I_R$ *(Homogeneous amplification),*
- $\boldsymbol{\alpha}_t$ *satisfies Assumptions 1.1,*

then under Assumptions A.2-A.5 of Gouriéroux et al. (2017), the C-ICA procedure, where the PML estimator of Gouriéroux et al. (2017) is employed in the ICA-step, is consistent.

Proof of Proposition 2.9. Let us consider the variance-covariance matrix of $\boldsymbol{u}_t$, when $T \to \infty$:

$$\Sigma_{\boldsymbol{u}} = ZZ' + H = (U\Sigma\Sigma'U' + \sigma_{\boldsymbol{e}}^2 I_N) = UDU', \text{ where } D = \Sigma\Sigma' + \sigma_{\boldsymbol{e}}^2 I_N.$$

We compute U by a spectral decomposition on $\Sigma_{\boldsymbol{u}}$ and we select the first R eigenvectors $U_R \in \mathbb{R}^{N\times R}$. We observe that

$$D = \left[\begin{array}{c|c} (\sigma^2 + \sigma_{\boldsymbol{e}}^2) I_R & 0 \\ \hline 0 & \sigma_{\boldsymbol{e}}^2 I_{N-R} \end{array}\right],$$

so we can estimate $\sigma_{\boldsymbol{e}}^2$ consistently in T by considering the last $N-R$ eigenvalues and we denote with $\widehat{H} = \hat{\sigma}_{\boldsymbol{e}}^2 I_N$ the estimate of the noise covariance matrix. Then, we estimate the amplification matrix with $\hat{\Sigma}_+^2 = \widehat{D}_R - \hat{\sigma}_{\boldsymbol{e}}^2 I_R$, where D_R contains the first R eigenvalues of D, and it holds that: $\hat{\Sigma}_+^2 \xrightarrow{\mathbb{P}} \Sigma_+^2$ as $T \to \infty$. Let $SS' = D_R$ and $\Pi_1 = S^{-1}U_R'$, so:

$$\begin{aligned} \boldsymbol{u}_t^1 &= \Pi_1 \boldsymbol{u}_t = S^{-1}U_R'(U\Sigma C'\boldsymbol{\alpha}_t + \boldsymbol{e}_t) \\ &= S^{-1}(\Sigma_+ C'\boldsymbol{\alpha}_t + U_R'\boldsymbol{e}_t) = C'S^{-1}(\Sigma_+\boldsymbol{\alpha}_t + \boldsymbol{e}_t^1) = C'\boldsymbol{\beta}_t, \end{aligned}$$

where $\boldsymbol{e}_t^1 = CU_R'\boldsymbol{e}_t \in \mathbb{R}^R$ and $\boldsymbol{\beta}_t = S^{-1}(\Sigma_+\boldsymbol{\alpha}_t + \boldsymbol{e}_t^1)$. Since S^{-1} and Σ_+ are multiple of the identity, C' commutes with them. The noise term $\boldsymbol{e}_t^1$ is Gaussian with variance-covariance matrix $\sigma_{\boldsymbol{e}}^2 I_R$. Then, $\text{cov}(\boldsymbol{\beta}_t) = S^{-1}(\sigma^2 I_R + \sigma_{\boldsymbol{e}}^2 I_R)S'^{-1} = I_R$. This means that $\boldsymbol{\beta}_t$ are i.i.d. with mutually independent components with at most one Gaussian component. Therefore, since $\boldsymbol{\beta}_t$ satisfies Assumptions 1.1, we can estimate the collapsed mixing matrix C' using standard ICA techniques. Moreover, since $\boldsymbol{e}_t^1$ is a Gaussian independent additive noise and independent from $\boldsymbol{\alpha}_t$, $\boldsymbol{\beta}_t$ satisfies Assumptions A.2-A.5 of Gouriéroux et al. (2017). Then, C' is estimated by considering the collapsed observations equation: $\boldsymbol{u}_t^1 = C'\boldsymbol{\beta}_t$. Finally we obtain an estimate of the mixing matrix with $\widehat{Z} = \widehat{U}_R\widehat{\Sigma}_+\widehat{C}'$, which is consistent since:

- $\hat{\Sigma}_+ \xrightarrow{\mathbb{P}} \Sigma_+$ as $T \to \infty$,
- $\widehat{C}' \xrightarrow{\mathbb{P}} C'$ as $T \to \infty$ by Gouriéroux et al. (2017),

and the product of consistent estimators is consistent, then $\widehat{Z} \xrightarrow{\mathbb{P}} Z$ as $T \to \infty$. The identifiability of U and C' is guaranteed by Lemma 2.5, indeed once we have obtained the first R columns of U the matrix C' is uniquely identified. □

Using PCA, when $T \to \infty$ the eigenvalues matrix of $\Sigma_{\boldsymbol{u}}$ will converge to a diagonal matrix, where exactly the first R eigenvalues will be equal to $\sigma^2 + \sigma_{\boldsymbol{e}}^2$ and the remaining $N-R$ to $\sigma_{\boldsymbol{e}}^2$ and the associated eigenvectors matrix is given by U. Therefore, asymptotically we can select R by analyzing the spectrum of $\Sigma_{\boldsymbol{u}}$. In our application we select R using PCA so that the cumulative percentage of variance of the first R components is greater than 85%.

Then, the first R eigenvectors U_R can be also conveniently used for the initialization step of Algorithm 1. Thus, from the last $N-R$ eigenvalues of $\boldsymbol{\Sigma_u}$ we estimate $\sigma^2_{\boldsymbol{e}}$, and let D_R be the diagonal matrix with the first R eigenvalues, the matrix $\boldsymbol{\Sigma}^2_+ = \sigma^2 I_R$ is estimated by considering $D_R - \sigma^2_{\boldsymbol{e}} I_R$. Therefore, the projection matrix is given by $\boldsymbol{\Pi}_1 = S^{-1}U'_R$, where S is the square root of D_R, i.e., $SS' = D_R$. Finally, once obtained $\widehat{C}'$ from the ICA-step, the matrix Z is estimated by $\widehat{Z} = \widehat{U}_R \widehat{\boldsymbol{\Sigma}}_+ \widehat{C}'$.

In other words, when the amplification matrix is homogeneous, i.e., $\boldsymbol{\Sigma}_+ = \sigma I_R$, the sensor noise formulation is equivalent to the source noise one and the collapsed mixing matrix C' can be estimated directly from the virtual sensors.

Remark 2.10. The projection matrix $\boldsymbol{\Pi}_1$ is related to the PCA estimates of the mixing matrix and the variance-covariance matrix of the noise. Indeed, when we select the first R eigenvalues and the associated eigenvectors of D, $\widehat{Z} = U_R S$, and $\widehat{H} = \widetilde{\sigma}^2_{\boldsymbol{e}} I_N$, and using the collapsing procedure we obtain that $\boldsymbol{\Pi}_1 = S^{-1}U'_R$, which is exactly the projection matrix used by the C-ICA procedure.

Remark 2.11. Once we have collapsed the signal we have not to standardize the virtual sensors $\boldsymbol{u}^1_t$. Indeed,

$$\begin{aligned}\mathbb{E}[\boldsymbol{u}^1_t \boldsymbol{u}^{1\prime}_t] &= \boldsymbol{\Pi}_1 \mathbb{E}[\boldsymbol{u}_t \boldsymbol{u}'_t] \boldsymbol{\Pi}'_1 = S^{-1}U'_R U D U' U_R S'^{-1} \\ &= S^{-1}\left[\ I_R \ \middle|\ 0\ \right] D \left[\frac{I_R}{0}\right] S^{-1} = S^{-1} D_R S'^{-1} = I_R.\end{aligned}$$

Remark 2.12. We now discuss the implication of relaxing the assumption of homogeneous amplification, i.e., $\boldsymbol{\Sigma}_+ = \sigma I_R$. If we assume an inhomogeneous amplification, $\boldsymbol{\Sigma}_+ = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_R)$, which can be consistently estimated with (the square root of) $\widehat{\boldsymbol{\Sigma}}^2_+ = \widehat{D}_R - \widehat{\sigma}^2_{\boldsymbol{e}} I_R$, then, the collapsed observations are given by $\boldsymbol{u}^1_t = S^{-1}(\boldsymbol{\Sigma}_+ C' \boldsymbol{\alpha}_t + U'_R \boldsymbol{e}_t)$. Contrary to the homogeneous case $\boldsymbol{\Sigma}_+$ does not commute with C'. Moreover, let us denote with $\boldsymbol{e}^1_t = C\boldsymbol{\Sigma}^{-1}_+ U'_R \boldsymbol{e}_t$ so that $\boldsymbol{u}^1_t = S^{-1}\boldsymbol{\Sigma}_+ C'(\boldsymbol{\alpha}_t + \boldsymbol{e}^1_t)$, where $\mathrm{Cov}(\boldsymbol{e}^1_t) = C'\boldsymbol{\Sigma}^{-1}_+ U'_R U_R \boldsymbol{\Sigma}^{-1}_+ C = C'\boldsymbol{\Sigma}^{-2}_+ C$. Without imposing other restrictions on C', this variance-covariance matrix is not diagonal. Then, since $\boldsymbol{e}^1_t$ is Gaussian, there is a dependence among the contaminated shocks $(\boldsymbol{\alpha}_t + \boldsymbol{e}^1_t)$ and Assumptions 1.1 are violated. So, we can not directly estimate C' using standard ICA techniques. However, we may still consider $\boldsymbol{u}^1_t$ as collapsed observations and implicitly assume that the problem $\boldsymbol{u}^1_t = G\boldsymbol{\beta}_t$, is well posed for some orthogonal matrix $G \in \mathbb{R}^{R\times R}$ and $\boldsymbol{\beta}_t \in \mathbb{R}^R$ such that $\boldsymbol{\Sigma}_{\boldsymbol{\beta}} = I_R$. The components of $\boldsymbol{\beta}_t$ are contaminated shocks since the collapsing procedure does not completely remove the noise from the structural shocks. So, we can identify the mixing matrix G by employing a standard ICA technique to the collapsed observations, which leads to the separation of the contaminated shocks $\boldsymbol{\beta}$ (not the pure one $\boldsymbol{\alpha}$), see Joho et al. (2000).

Summarizing, when the noise is Gaussian, independent and homoskedastic, we can

estimate Z even if N is bounded. Indeed, by collapsing all the information in R collapsed observations, we can employ classical ICA techniques on these observations to get a T consistent estimator of C, under the assumption of homogeneous amplification. Moreover, the inhomogeneous amplification case is closely related to the heteroskedastic noise, i.e., when the noise is not homoskedastic. If $\Sigma_+ = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_R)$ and $H = \sigma_e^2 I_N$, then, by employing the previous collapsing procedure, we obtain $\boldsymbol{u}_t^1 = S^{-1}(\Sigma_+ C'\boldsymbol{\alpha}_t + U_R'\boldsymbol{e}_t) = S^{-1}\Sigma_+(C'\boldsymbol{\alpha}_t + \Sigma_+^{-1}U_R'\boldsymbol{e}_t)$, so that

$$\Sigma_+^{-1} S \boldsymbol{u}_t^1 = C'\boldsymbol{\alpha}_t + \boldsymbol{e}_t^1, \tag{2.6}$$

where $\boldsymbol{e}_t^1 = \Sigma_+^{-1}U_R'\boldsymbol{e}_t$ with variance-covariance $\Sigma_{\boldsymbol{e}^1} = \sigma_{\boldsymbol{e}}^2 \cdot (\Sigma_+\Sigma_+)^{-1}$. Thus, the inhomogeneous amplification implies an heteroskedastic noise in the collapsed equation (2.6). Since, problem (2.6) is now a non-singular square problem one could use the algorithm of Hyvärinen (1999) without the need to ("quasi") whitening the data, Hyvärinen et al. (2001).

2.2.3.2 The heteroskedastic noise case

In the most general case we have $\boldsymbol{u}_t = Z\boldsymbol{\alpha}_t + \boldsymbol{e}_t$, where $\boldsymbol{u}_t, \boldsymbol{e}_t \in \mathbb{R}^N$, $\boldsymbol{\alpha}_t \in \mathbb{R}^R$, $Z \in \mathbb{R}^{N\times R}$ and $H = \text{Cov}(\boldsymbol{e}_t)$ can be a diagonal matrix with different elements or a generic matrix with non-zero covariance elements. Unfortunately in this case we can not disentangle the noise from the structural shocks mix $Z\boldsymbol{\alpha}_t$ without require that both N and T go to infinity. Therefore, this case would pose a problem to the standard VAR approach (where N should be bounded). However, for completeness, we propose a general approach which combines both DFM approach and ICA technique, where the noisy case is reduced to the noiseless one discussed in Section 2.2.2.

Let us consider the static representation of a DFM,

$$\boldsymbol{y}_t = \Lambda \boldsymbol{f}_t + \boldsymbol{e}_t, \tag{2.7}$$

$$\boldsymbol{f}_t = B(L)\boldsymbol{v}_t, \tag{2.8}$$

where $\boldsymbol{y}_t \in \mathbb{R}^M$ are the observed macroeconomic variables, $\boldsymbol{e}_t \in \mathbb{R}^M$ are the idiosyncratic noise, e.g. the measurement errors of $\boldsymbol{y}_t$, $\boldsymbol{f}_t \in \mathbb{R}^N$ are the common static factors, $\Lambda \in \mathbb{R}^{M\times N}$ are the factor loadings, $\boldsymbol{v}_t \in \mathbb{R}^R$, $B(L) \in \mathbb{R}^{N\times R}$ is a matrix polynomial, L is the lag operator and generally $M > N \geq R$. Anderson and Deistler (2008b) have shown that in the singular case when $N > R$ the polynomial $B(L)$ is invertible, so $\boldsymbol{v}_t$ are *fundamental* shocks, and the inverse is a finite polynomial, see Anderson and Deistler (2008a,b), and Forni et al. (2009) for a complete discussion. This implies that there exists an equivalent finite (singular) VAR representation for the common factor Eq. (2.8), i.e., $A(L)\boldsymbol{f}_t = B(0)\boldsymbol{v}_t$, where $\boldsymbol{u}_t = B(0)\boldsymbol{v}_t \in R^N$ are the reduced shocks of the above VAR

which are related to corresponding structural shocks $\boldsymbol{\alpha}_t \in \mathbb{R}^R$ as $\boldsymbol{u}_t = Z\boldsymbol{\alpha}_t$, where $Z \in \mathbb{R}^{N\times R}$. Then, thanks to the DFM approach, we have reduced the most generic (heteroskedastic) noisy case to the noiseless case, indeed, when M goes to infinity the idiosyncratic component can be separated from the common factors, which leads to the standard noiseless case discussed in Section 2.2.2.

In this case we may employ a combination of the recent approach of Doz et al. (2012) for the estimation of (dynamic) factor models, together with the C-ICA procedure directly on the reduced shocks. However, the resulting procedure is not feasible for the VAR approach, since it requires that both and N and T must go to infinity. For completeness, we report in Appendix A.1 this approach.

2.3 Empirical analysis

In our empirical analysis we follow Gouriéroux et al. (2017) considering a VAR model with the following four macroeconomic variables: inflation, π, unemployment gap, U_g, output gap, O_g, and the nominal short-term interest rate, r. The sample period considered is 1959:Q1−2015:Q1. All data are collected from the Federal Reserve Economic Database (FRED) at quarterly frequency. Inflation is computed as the change in the logarithm of the GDP deflator. The unemployment gap is obtained as the difference between the observed unemployment rate (FRED mnemonic UNRATE) and the natural rate of unemployment (FRED mnemonic NROU). The output gap is calculated as the variation of the natural logarithm of real GDP (FRED mnemonic GDPC1) from the log potential GDP (FRED mnemonic GDPPOT). The nominal short-term interest rate is given by the federal funds rate (FRED mnemonic FEDFUNDS).

Both unemployment and output gap are proxies for the economic activity. Gouriéroux et al. (2017) have found that the scheme $[\pi, U_g, r]$ for the short-run (Cholesky) restrictions has to be preferred to the one containing the output gap, as it is not rejected by their test. However, our aim is to combine the information coming from both variables in the identification of the economic activity shocks and in the construction of the corresponding IRFs. This goal can be achieved by considering a singular SVAR having more macroeconomic series than structural shocks. Then, contrary to Gouriéroux et al. (2017) we consider the following scheme: $[\pi, U_g, O_g, r]$. So, we have $N = 4$ and $T = 224$ number of observations. We estimate a 6-lag VAR with intercept term and an exogenous variable given by the variation in the logarithm of oil prices[8]. Performing a principal component analysis on the residual of the VAR, we found that the three principal components explain about 88% of the total variance (where each of the first three principal components explain more than the 20% of the total variance). Accordingly, we set the number of the structural shocks equal to $R = 3$, which is in agreement with Gouriéroux et al. (2017),

[8]In order to replicate the results of Gouriéroux et al. (2017) we estimate exactly the same VAR model.

where they consider only three macroeconomic variables for estimating three structural shocks.

We account for possible parameter and model uncertainty in the reduced form VAR by allowing the presence of noise in the observations. In particular, we assume that the variance-covariance matrix of the noise H is homoskedastic and an homogeneous amplification for the mixing matrix. Then, different IRFs are computed according to the following identification schemes: i) the PML-ICA on 4 structural shocks, ii) the short-run restrictions on 4 structural shocks and iii) the C-ICA approach on 3 structural shocks. For the PML estimator we consider asymmetric mixtures of Gaussian distributions as pseudo density functions with the same specification[9] of Gouriéroux et al. (2017). The short-run restrictions are obtained using the Cholesky decomposition of the variance-covariance matrix of the reduced residuals. Then, with the first two approaches we obtain 4 IRFs for 4 structural shocks while for the C-ICA there are 4 IRFs related to 3 structural shocks (demand, monetary policy and supply shocks).

Before proceeding with the estimation of the IRFs we investigate the statistical properties of the residuals. We find no evidence of statistically significant dependency on time and Jarque-Bera tests reject the null hypothesis of Gaussian distribution for all the four residuals. We also test the autocorrelation and normality for the collapsed observations which are used for the C-ICA procedure and we find analogous results, i.e., no significant autocorrelation and clear rejection of gaussianity assumptions.

For the short-run restrictions the confidence interval at 90% level from 100 bootstrap samples on the residuals is computed and the corresponding median is reported. The IRFs are interpreted using the same economic indications of Gouriéroux et al. (2017). According to basic economic theory a shock is labelled as contractionary monetary-policy shock if it has a short and medium term positive impact on interest rate and a negative impact on both inflation and economic activity. The supply (demand) shock is expected to affect in the short and medium term economic activity and inflation with opposite (same) signs. When the response on output gap is negative (positive) the shock is labelled as contractionary (expansionary).

Figure 2.1 shows the IRF of a contractionary supply shock, of an expansionary demand shock and of an contractionary monetary policy shock, respectively. For the C-ICA

[9] We select three distinct asymmetric mixtures of Gaussian distributions as pseudo density functions. Each of the pseudo-density corresponds to the density of a random variable equal to $X_i = B_i Y_{i,1} + (1 - B_i)Y_{i,2}$ where B_i is a Bernoulli random variable of parameter p_i, $Y_{i,1} \sim N(\mu_{i,1}, \sigma^2_{i,1})$ and $Y_{i,2} \sim N(\mu_{i,2}, \sigma^2_{i,2})$. The expectation and variance of each Gaussian mixture X_i are respectively equal to zero and one. The parameters of the Bernoulli random variables are set to $p_1 = p_2 = p_3 = 0.5$. The mean and standard deviation of $Y_{i,1}$ for $i = 1, 2, 3$ are set to $\mu_{1,1} = \mu_{2,1} = \mu_{3,1} = 0.1$ and $\sigma_{1,1} = 0.5$, $\sigma_{2,1} = 0.7$ and $\sigma_{3,1} = 1.3$, so that $\mu_{1,2} = \mu_{2,2} = \mu_{3,2} = -0.1$ and $\sigma_{1,2} = 1.32$, $\sigma_{2,2} = 1.22$ and $\sigma_{3,2} = 0.54$. For further details see footnote 23 of Gouriéroux et al. (2017). For the case i) the fourth Gaussian mixture has $p_4 = 0.5$ for the Bernoulli parameter and $\mu_{4,1} = 0$, $\sigma_{4,1} = 0.9$, $\mu_{4,2} = 0$ and $\sigma_{4,2} = 1.08$ for the expectation and standard deviation of $Y_{4,1}$ and $Y_{4,2}$, respectively.

procedure the 90% level bootstrap confidence interval (from 100 bootstrap samples) is also reported[10] (dotted red lines). We note that the IRFs obtained with the C-ICA approach are in line with those of the short-run and PML.

Principal component analysis and the evidence in Gouriéroux et al. (2017) suggest that the system is most likely driven by only 3 structural shocks and thanks to the C-ICA procedure, we can use the entire set of information contained in both the unemployment and output gap to identify the structural shocks related to economic activity. The standard methods, however, by implicitly assuming 4 structural shocks would likely be misspecified. Figure 2.2 illustrates the IRF for the $4th$ shock obtained with the short-run restrictions and the PML approach. The economic interpretation of the reported IRF turns out to be difficult. Indeed, even if there is a negative impact on the interest rate, the effect on real activity and inflation are also negative, which means that we can not interpret this shock as a expansionary monetary policy shock. Moreover, the impact on inflation does not seem to be significantly different from zero, so that it is not clear whether this shock is related to supply or demand. In Appendix A.3 we report the estimated mixing matrices $\widehat{Z}$, the estimated orthogonal matrix $\widehat{C}'$ and the related statistics. The matrix C' estimated with the C-ICA procedure has almost all statistically significant coefficients, in contrast to that obtained with the standard PML-ICA.

In this last part of the empirical analysis we compare the IRFs obtained with the C-ICA approach on the scheme $[\pi, U_g, O_g, r]$ with those obtained by Gouriéroux et al. (2017) on the scheme $[\pi, U_g, r]$. Then, the same 6-lag VAR is estimated as before and the IRFs are obtained with the short-run restrictions and PML approach. We compare these IRFs with those obtained by using the C-ICA method in Figure 2.3. Again, Jarque-Bera tests reject the normal assumptions of the residuals and we find no evidence of statistically significant autocorrelation. The IRFs of C-ICA seem to be in line with those obtained with the PML approach, especially for the demand and monetary policy shocks. However, we stress that with the short-run restrictions and the PML approach, we use only the information of inflation, unemployment gap and interest rate, while the C-ICA procedure includes also the output gap in the identification of the 3 structural shocks, thus merging the information from different proxies. This, in turn, should translate in a more precise estimation of the IRFs. In Appendix A.3 we also report the estimated mixing matrices obtained when short-run restrictions and PML-ICA approach are employed on the scheme $[\pi, U_g, r]$. Furthermore, thanks to the C-ICA approach, also the IRF of the

[10] We use bootstrap on residuals, so that from each bootstrapped trajectory k we re-estimate the underlying reduced VAR model. Then, we obtain the corresponding $\widehat{Z}_k$ and we apply the permutation and change of signs in a such way that the product between the original (obtained with non-bootstrap data) estimate $\widehat{Z}'$ and the permuted $\widehat{Z}_k$ is proportional to the identity. Althought this product is not exactly proportional to the identity, we choose the permutation and change of signs so that the leading terms of columns are in the diagonal and they are positive, like in the identification scheme of Lanne et al. (2017). Therefore, for each step k the structural shocks have the same labelling and the relative IRFs may be compared.

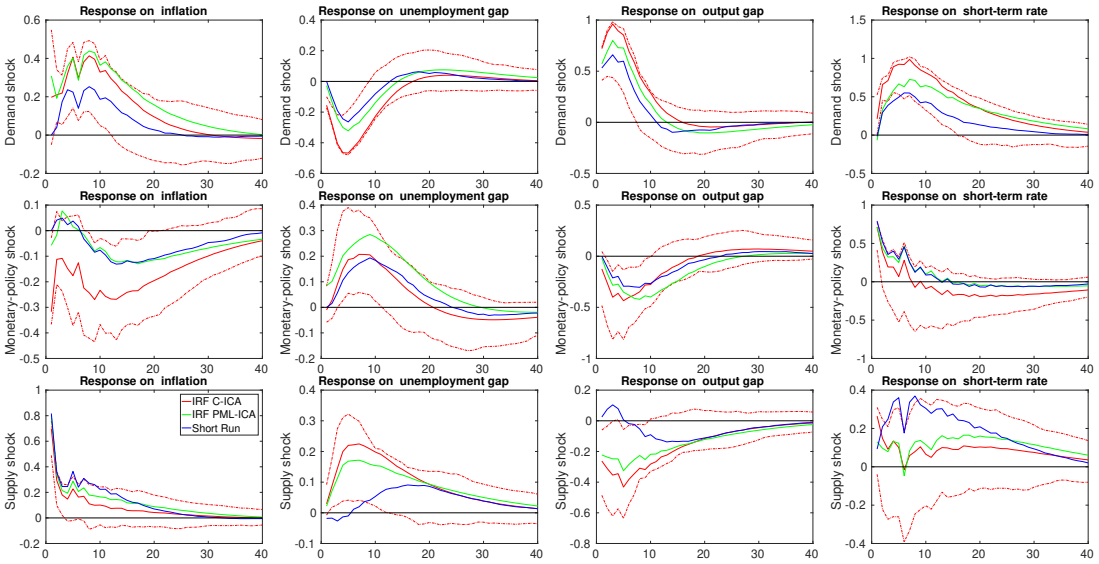


Fig. 2.1 Impulse response functions. This figure exhibits IRFs obtained from a VAR model on the scheme $[\pi, U_g, O_g, r]$ related to an expansionary demand shock, contractionary monetary policy shock and contractionary supply shock, respectively. The shocks are contractionary when the response on output gap is negative and expansionary when it is positive. The blue and green lines correspond to the IRFs derived by short-run (Cholesky) restrictions and the PML approach, respectively. Red lines correspond to the IRFs computed by the C-ICA procedure on an singular SVAR with 3 structural shocks. Dotted red lines represent the confidence interval at 90% level using the bootstrap method derived from C-ICA procedure over 100 bootstrap samples.

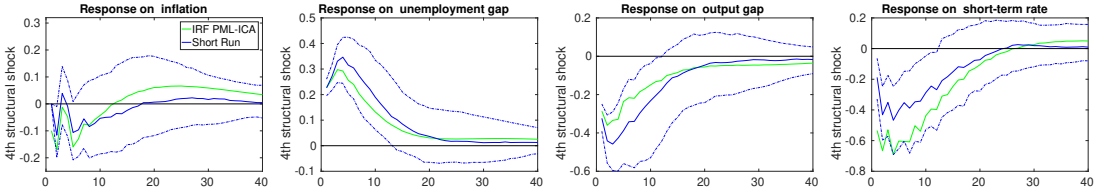


Fig. 2.2 Impulse response functions. This figure exhibits IRFs obtained from a VAR model on the scheme $[\pi, U_g, O_g, r]$ related to a $4th$ structural shock. The blue and green lines correspond to the IRFs derived by short-run (Cholesky) restrictions and the PML approach, respectively. Dotted blue lines represent the confidence interval at 90% level using the bootstrap method derived from the short-run restrictions over 100 bootstrap samples.

output gap can be constructed along with the other responses, which allows to obtain a more complete picture of the effects of the structural shocks on all the macroeconomic variables.

Finally, we investigate the variance decomposition related to the structural shocks. We denote with $4SR$ ($3SR$) and $4PML\text{-}ICA$ ($3PML\text{-}ICA$) the variance decomposition obtained by the 4-VAR (3-VAR) model using short-run restrictions and PML-ICA, respectively, and with $C\text{-}ICA$ the variance decomposition related to the C-ICA procedure,

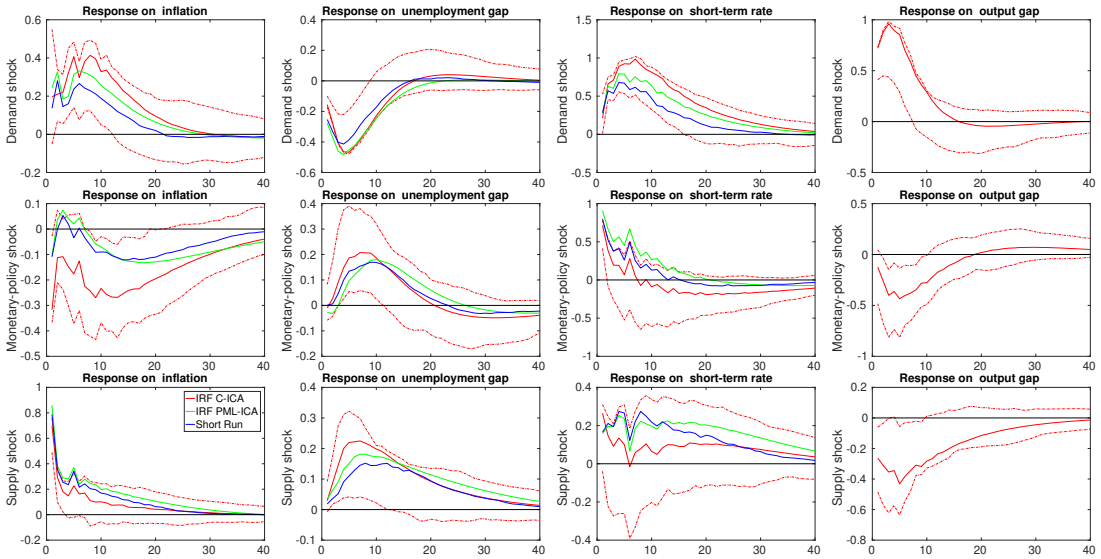


Fig. 2.3 Impulse response functions. This figure exhibits IRFs related to an expansionary demand shock, contractionary monetary policy shock and contractionary supply shock, respectively. The shocks are contractionary when the response on output gap is negative and expansionary when it is positive. The blue and green lines exhibits impulse response functions obtained from a VAR model on the scheme $[\pi, U_g, r]$ which are derived by short-run (Cholesky) restrictions and the PML approach, respectively. The red lines correspond to the IRFs computed by the C-ICA procedure on a singular SVAR model with 3 structural shocks on the scheme $[\pi, U_g, O_g, r]$. Dotted red lines represent the confidence interval at 90% level using the bootstrap method derived from the C-ICA procedure over 100 bootstrap samples.

see Table 2.1, 2.2 and 2.3.[11] Overall, we observe that the C-ICA exhibits a more uniform variance decomposition among the variables compared to the other methods. Indeed, even if the monetary policy long-run effect on the real economy is found by all the identification techniques, the effect on the inflation is observed only for the C-ICA procedure. In particular, using the C-ICA procedure we found that after four years the monetary policy shock explains about 27% of the variance of inflation. Moreover, while for the standard identification methods monetary policy shock does not seem to explain the inflation even in the short-run, this effect is found by the C-ICA procedure. Therefore, in the C-ICA identification, the variance of inflation is explained by monetary policy shock both in the long-run and short-run, which can be arguably considered to be more in accordance with macroeconomic theory.

We conclude that using the C-ICA procedure we are able to conduct a more accurate IRFs analysis and variance decomposition of the macroeconomic variables in the system.. The IRFs analysis and the variance decomposition confirm that C-ICA correctly merges the information contained in all the variables thus allowing a more precise and coherent identification of the low dimensional system of structural shocks driving the economy.

[11] In Appendix A.3 the variance decomposition of the 4-th structural shocks is reported for completeness.

Table 2.1 Variance Decomposition related to demand shock. For each VAR model considered the variance explained by the demand shock relative to each variables is reported in columns. The lag period after the shock is measured in quarters. Values in parentheses are the standard deviations computed from 100 bootstrap samples.

	Demand Shock									
Lag	0		2		4		8		16	
$4SR$										
π	0	(0)	0.0354	(0.0290)	0.1136	(0.0522)	0.1900	(0.0841)	0.2165	(0.1054)
U_g	0	(0)	0.1608	(0.0514)	0.2639	(0.0841)	0.2539	(0.1069)	0.2080	(0.0976)
O_g	0.7281	(0.0579)	0.6482	(0.0778)	0.6293	(0.0946)	0.5428	(0.1172)	0.4790	(0.1184)
r	0.0004	(0.0087)	0.1377	(0.0542)	0.224	(0.0754)	0.3231	(0.1079)	0.3504	(0.1279)
$4PML\text{-}ICA$										
π	0.1267	(0.0383)	0.1951	(0.0398)	0.3315	(0.0425)	0.4841	(0.0541)	0.5640	(0.0924)
U_g	0.0123	(0.0545)	0.2229	(0.0562)	0.3250	(0.0784)	0.3144	(0.1182)	0.2425	(0.1465)
O_g	0.7123	(0.0448)	0.7073	(0.0641)	0.6896	(0.0795)	0.5951	(0.1121)	0.5022	(0.1301)
r	0.0055	(0.1312)	0.1543	(0.1248)	0.2792	(0.1092)	0.3932	(0.0991)	0.4839	(0.0864)
$C\text{-}ICA$										
π	0.0620	(0.1529)	0.1520	(0.1608)	0.3178	(0.1765)	0.4590	(0.1968)	0.4712	(0.2061)
U_g	0.9619	(0.1628)	0.8731	(0.1857)	0.7924	(0.1988)	0.6951	(0.2030)	0.6061	(0.1855)
O_g	0.8600	(0.1977)	0.8051	(0.2085)	0.7630	(0.2112)	0.6970	(0.2071)	0.6508	(0.1959)
r	0.0716	(0.1832)	0.5520	(0.1440)	0.7521	(0.1186)	0.8603	(0.1245)	0.8869	(0.1419)
$3SR$										
π	0	(0)	0.0510	(0.0323)	0.0767	(0.0434)	0.1555	(0.0768)	0.1651	(0.0980)
U_g	0.9810	(0.0224)	0.9846	(0.0222)	0.9542	(0.0405)	0.8683	(0.0870)	0.7490	(0.1294)
r	0.1178	(0.0606)	0.3446	(0.0967)	0.4561	(0.1031)	0.5260	(0.1240)	0.5448	(0.1427)
$3PML\text{-}ICA$										
π	0.0735	(0.1644)	0.1692	(0.1716)	0.2192	(0.1813)	0.3342	(0.1938)	0.3599	(0.1984)
U_g	0.9749	(0.1254)	0.9656	(0.1343)	0.9275	(0.1492)	0.8291	(0.1751)	0.7026	(0.1806)
r	0.0830	(0.1316)	0.3390	(0.1658)	0.4788	(0.1708)	0.5550	(0.1784)	0.6053	(0.1842)

Table 2.2 Variance Decomposition related to monetary policy shock. For each VAR model considered the variance explained by the monetary policy shock relative to each variables is reported in columns. The lag period after the shock is measured in quarters. Values in parentheses are the standard deviations computed from 100 bootstrap samples.

	Monetary Policy Shock									
Lag	0		2		4		8		16	
$4SR$										
π	0	(0)	0.0046	(0.0125)	0.0050	(0.0180)	0.0102	(0.0212)	0.0556	(0.0644)
U_g	0	(0)	0.0121	(0.0114)	0.0497	(0.0353)	0.1317	(0.0817)	0.2064	(0.1260)
O_g	0	(0)	0.0341	(0.0217)	0.0643	(0.0419)	0.1328	(0.0847)	0.1721	(0.1131)
r	0.9345	(0.0518)	0.5959	(0.0951)	0.4191	(0.0991)	0.3018	(0.1023)	0.2305	(0.0958)
$4PML\text{-}ICA$										
π	0.0044	(0.0383)	0.0090	(0.0398)	0.0084	(0.0425)	0.0091	(0.0541)	0.0356	(0.0924)
U_g	0.1203	(0.0545)	0.1166	(0.0562)	0.1584	(0.0784)	0.2580	(0.1182)	0.3614	(0.1465)
O_g	0.0008	(0.0448)	0.0533	(0.0641)	0.0845	(0.0795)	0.1757	(0.1121)	0.2517	(0.1301)
r	0.6235	(0.1312)	0.3692	(0.1248)	0.2437	(0.1092)	0.1748	(0.0991)	0.1245	(0.0864)
$C\text{-}ICA$										
π	0.1606	(0.0797)	0.1450	(0.0893)	0.1417	(0.1157)	0.1796	(0.1589)	0.2758	(0.1856)
U_g	0.0039	(0.0968)	0.0297	(0.1527)	0.0744	(0.1825)	0.1305	(0.2018)	0.1611	(0.1901)
O_g	0.0259	(0.1636)	0.0890	(0.1963)	0.1154	(0.2049)	0.1525	(0.2086)	0.1638	(0.1975)
r	0.8166	(0.1999)	0.3917	(0.1407)	0.2112	(0.1004)	0.1183	(0.1001)	0.093	(0.1194)
$3SR$										
π	0	(0)	0.0118	(0.0174)	0.0131	(0.0223)	0.0116	(0.0226)	0.0379	(0.0465)
U_g	0	(0)	0.0057	(0.0089)	0.0409	(0.0340)	0.1213	(0.0791)	0.2094	(0.1247)
r	0.8665	(0.0669)	0.5830	(0.1028)	0.4335	(0.1058)	0.3557	(0.1148)	0.2884	(0.1151)
$3PML\text{-}ICA$										
π	0.0134	(0.0898)	0.0149	(0.0762)	0.0128	(0.0704)	0.0117	(0.0714)	0.0448	(0.0945)
U_g	0.0098	(0.0500)	0.0041	(0.0490)	0.0150	(0.0650)	0.0632	(0.1010)	0.1282	(0.1271)
r	0.8901	(0.1509)	0.6198	(0.1763)	0.4704	(0.1704)	0.3971	(0.1693)	0.3232	(0.1557)

Table 2.3 Variance Decomposition related to supply shock. For each VAR model considered the variance explained by the supply shock relative to each variables is reported in columns. The lag period after the shock is measured in quarters. Values in parentheses are the standard deviations computed from 100 bootstrap samples.

	Supply Shock									
Lag	0		2		4		8		16	
$4SR$										
π	1	(0)	0.9435	(0.0317)	0.8598	(0.0548)	0.7705	(0.0865)	0.7004	(0.1120)
U_g	0.0064	(0.0170)	0.0047	(0.0182)	0.0025	(0.0172)	0.0043	(0.0257)	0.0357	(0.0623)
O_g	0.0016	(0.0129)	0.0103	(0.0209)	0.0083	(0.0219)	0.0097	(0.0281)	0.0368	(0.0559)
r	0.0128	(0.0198)	0.0662	(0.0436)	0.1181	(0.0607)	0.1408	(0.0775)	0.1938	(0.1073)
$4PML\text{-}ICA$										
π	0.8557	(0.0383)	0.7594	(0.0398)	0.6159	(0.0425)	0.4645	(0.0541)	0.3689	(0.0924)
U_g	0.0080	(0.0545)	0.0533	(0.0562)	0.0849	(0.0784)	0.1118	(0.1182)	0.1437	(0.1465)
O_g	0.1080	(0.0448)	0.0815	(0.0641)	0.0920	(0.0795)	0.1100	(0.1121)	0.1334	(0.1301)
r	0.0208	(0.1312)	0.0149	(0.1248)	0.0162	(0.1092)	0.0147	(0.0991)	0.0250	(0.0864)
$C\text{-}ICA$										
π	0.7774	(0.1614)	0.7030	(0.1611)	0.5406	(0.1521)	0.3614	(0.1442)	0.2530	(0.1490)
U_g	0.0342	(0.1168)	0.0972	(0.1323)	0.1332	(0.1418)	0.1745	(0.1494)	0.2327	(0.1363)
O_g	0.1141	(0.1399)	0.1059	(0.1354)	0.1216	(0.1406)	0.1505	(0.1432)	0.1854	(0.1342)
r	0.1118	(0.0856)	0.0563	(0.0710)	0.0367	(0.0739)	0.0214	(0.0777)	0.0201	(0.0968)
$3SR$										
π	1	(0)	0.9372	(0.0313)	0.9102	(0.0456)	0.8329	(0.0788)	0.7970	(0.1021)
U_g	0.0190	(0.0224)	0.0097	(0.0204)	0.0049	(0.0196)	0.0105	(0.0346)	0.0417	(0.0695)
r	0.0158	(0.0222)	0.0723	(0.0465)	0.1104	(0.0631)	0.1183	(0.0770)	0.1667	(0.1158)
$3PML\text{-}ICA$										
π	0.9130	(0.1714)	0.8159	(0.1774)	0.7681	(0.1877)	0.6540	(0.1979)	0.5953	(0.1967)
U_g	0.0153	(0.1114)	0.0303	(0.1228)	0.0574	(0.1425)	0.1077	(0.1644)	0.1692	(0.1622)
r	0.0269	(0.0820)	0.0412	(0.1122)	0.0508	(0.1337)	0.0479	(0.1428)	0.0715	(0.1647)

Appendix A

Appendix of Chapter 2

A.1 A DFM-ICA procedure

We consider the quasi-maximum likelihood estimator of Doz et al. (2012) which guaranteed consistency of the estimator of the sources even in presence of misspecification of the true distribution. In the first step, we assume a misspecification of the true sources and we can employ the EM-procedure using the Kalman algorithm. Then, being $R < N$, the Kalman smoother step can be improved using the collapsing procedure, as it is observed by Doz et al. (2012). Finally, by exploiting the asymptotic results of Doz et al. (2012) we can obtain a consistent estimator of the sources, $\widehat{\boldsymbol{\alpha}}_t$. In particular, the bias induced by the assumed misspecification tends to zero as both N and T go to infinity. However, the sources remain identified only up to a rotation which can be estimated with ICA on the estimated sources, i.e, we separate the sources by employing ICA directly on $\widehat{\boldsymbol{\alpha}}_t$, so that we obtain the QML-Collapsing-ICA approach. We observe that in this case the collapsing procedure acts as a computational tool to improve the efficiency of the numerical computations of the Kalman smoother, (see Jungbacker and Koopman, 2008, for further details). We will then refer to the previous method as the Q-CICA procedure, see Algorithm 2.

Algorithm 2 The Q-CICA procedure.

Assuming a misspecification of the true sources:

1: *QML step*: Perform the EM-procedure recursively using Kalman algorithm and get an estimate of the system matrices Z and H. Improve the Kalman smoother step with the collapsing procedure and obtain the estimate of the sources.
2: *ICA-Step*: Separate the sources employing the ICA directly on $\widehat{\boldsymbol{\alpha}}_t$.

Use the identified mixing matrix Z for the IRF analysis.

Proposition A.1. *Given*

$$\boldsymbol{u}_t = Z\alpha_t + \boldsymbol{e}_t, \text{ where } Z \in \mathbb{R}^{N\times R} \text{ and } rank(Z) = R, \quad \boldsymbol{u}_t, \boldsymbol{e}_t \in \mathbb{R}^N, \boldsymbol{\alpha} \in \mathbb{R}^R.$$

Let $Z = U\Sigma C'$ a singular value decomposition of Z, where $U \in \mathbb{R}^{N\times N}, C' \in \mathbb{R}^{R\times R}$ are orthogonal matrices and $\Sigma = \begin{bmatrix} \Sigma_+ \\ 0 \end{bmatrix}$. Let us suppose the assumptions of Doz et al. (2012) are satisfied and:

- *$\boldsymbol{e}_t \sim N(0, H)$, H where $\boldsymbol{e}_t$ is independent from $\boldsymbol{\alpha}_t$ and H is a generic variance-covariance matrix with possible cross covariance terms (i.e. as for an approximate factor model),*
- *$\boldsymbol{\alpha}_t$ satisfies Assumptions 1.1,*

then the Q-CICA procedure is consistent for the identification of the sources when N and T go to infinity.

Proof. By Doz et al. (2012) we obtain an estimate of the sources. We remark that we can estimate the space of the common factors (sources), i.e., if $\widehat{Z}$, $\widehat{\boldsymbol{\alpha}}_t$ denote the estimate of Z and $\boldsymbol{\alpha}_t$, then

$$\widehat{Z}\widehat{\boldsymbol{\alpha}}_t \xrightarrow{\mathbb{P}} Z\boldsymbol{\alpha}_t \text{ as } N, T \to \infty,$$

by Doz et al. (2012). In other words, we can disentangle the noise from the sources, but

$$\widehat{Z} \xrightarrow{\mathbb{P}} Z \cdot C, \quad \widehat{\boldsymbol{\alpha}}_t \xrightarrow{\mathbb{P}} C' \cdot \boldsymbol{\alpha}_t, \qquad \text{as } N, T \to \infty,$$

where $CC' = C'C = I_R$. This mean that $\widehat{\boldsymbol{\alpha}}_t$ is a rotation of the consistent estimator of the true sources. However, we can estimate C using ICA on the estimated sources and we obtain $\widehat{C} \xrightarrow{\mathbb{P}} C$, as $T \to \infty$. Finally we get an estimate of Z by considering

$$\widehat{Z}\widehat{C}' \xrightarrow{\mathbb{P}} Z \text{ as } N, T \to \infty.$$

□

A.2 Simulation analysis

In the following simulation analysis we will investigate the convergence and the finite sample properties of the proposed C-ICA procedure. The simulation setting will assume standardized sources and homoskedastic noise with an homogeneous amplification matrix, $\sigma^2_{\boldsymbol{\alpha}} = 1$, $H = \sigma^2_{\boldsymbol{e}} I_N$, $\Sigma_+ = I_R$ and $R = 3$ structural shocks. The sources $\boldsymbol{\alpha}$ are drawn from a Student t-distribution, with different degrees of freedom $\nu = 10, 15, 5$. At each step of the simulation we generate the noise term $\boldsymbol{e}_t$ from a Gaussian distribution, as assumed so far, and we recover the related observations $\boldsymbol{u}_t$ with a fixed mixing matrix Z. We analyze the finite sample properties of the collapsing procedure in separating the sources in a realistic parameters setting, i.e., with a number of observations and variables close to those encounter in the empirical applications. For the ICA step we employ the same specification of Gouriéroux et al. (2017) for their PML estimator.

Thus, for each simulation k we compute $\widehat{U}_{R,k}$ and the amplification matrix $\widehat{\Sigma}_{+,k}$. We estimate the collapsed mixing matrix $\widehat{C}'_k$ using the ICA technique, so that for each simulated path we obtain an estimate of the mixing matrix $\widehat{Z}_k = \widehat{U}_{R,k}\widehat{\Sigma}_{+,k}\widehat{C}'_k$.

We report both numerical and graphical results of the finite sample procedure in separating the shocks. We perform a sensitivity analysis by varying noise variance $\sigma^2_{\boldsymbol{e}} = 0.10, 0.25, 0.50$, the sample size $T = 100, 200, 2000$ and the number of variables $N = 5, 10, 20$. For each parameter setting $N_{sim} = 100$ simulations are performed. We refer to $\widehat{\boldsymbol{\alpha}}_{t,k} = \widehat{Z}'_k \cdot Z\boldsymbol{\alpha}_t$ as the unmixed shocks and to $\boldsymbol{\beta}_{t,k} = \widehat{Z}'_k \boldsymbol{u}^1_{t,k}$ as the contaminated shocks. For each simulation the signal-to-noise ratio (SNR) is computed. For a general signal $\boldsymbol{x}_t$ we denote by $SNR_{\boldsymbol{x}}$ the SNR of $\boldsymbol{x}$ over the noise, which is defined as the ratio between the variance of the signal $\boldsymbol{x}_t$ and that of the noise term $\boldsymbol{e}_t$. We denote by sys_t the signal provided by the theoretical signal $\boldsymbol{\alpha}_t + \boldsymbol{e}_t$. In Tables A.1 and A.2 we report the average SNR (among the simulations) of the observations $\boldsymbol{u}_t$, of the collapsed observations $\boldsymbol{u}^1_t$ and of $\boldsymbol{\beta}$ together with the average estimates of σ and $\sigma^2_{\boldsymbol{e}}$.

Regardless of parameter settings, it is clear that the shocks signal on the original observations is very poor, being essentially masked by the noise. In contrast, by looking at the SNR on the collapsed observations, it appears that the collapsing step is able to reduce the noise, allowing the ICA procedure to separate the structural shocks even if it is applied on the collapsed variables. Indeed, a marked improvement of the SNR is observed after the application of collapsing procedure. In particular, even when $N = 5$, i.e. N is close to the number of structural shocks, the collapsing procedure provides the same SNR improvement observed for $N = 10$. The $SNR_{\boldsymbol{u}^1_t}$ becomes close to the theoretical one, SNR_{sys}, indicating that applying the ICA on the collapsed observations considerably enhances the accuracy of the separation of the shocks. Hence, these results show that applying the collapsing procedure has a powerful denoising effect. Moreover, the estimates of σ and $\sigma^2_{\boldsymbol{e}}$ turn out to be T-consistent regardless of N.

We investigate the convergence of the C-ICA procedure by computing the identification scheme of Lanne et al. (2017) on $\widehat{Z}'_k Z$ to obtain $\widehat{A}_k = ID(\widehat{Z}'_k Z)$, where ID denotes the identification scheme operator. Indeed, under general assumptions ICA identifies Z in an equivalence class in the set of nonsingular matrices. If $\sim$ is the equivalence relation w.r.t. permutations and scalings, then ICA estimates the equivalence class of Z. However, it is possible to select a particular representative by using the identification scheme of Lanne et al. (2017), so that if $A \sim B$, then $ID(A) = ID(B)$, see Lanne et al. (2017) for further details. We use this result for our simulations analysis. Therefore, given two representatives A and B of two equivalence classes we want to verify whether the two equivalence classes are the same, i.e., $ID(A) = ID(B)$. For this purpose we apply ID on A and B. If the mixing matrix were correctly estimated we should observe $ID(\widehat{Z}'_k Z) = \widehat{A}_k = I_R$. Thus, we consider as penalizing function the mean squared error w.r.t. the Frobenius norm, i.e., MSE1 $= \frac{1}{N_{sim}} \sum_{i=1}^{N_{sim}} ||\widehat{A}_k - I_R||^2_F$. Furthermore, we com-

Table A.1 The average SNR of the observations $\boldsymbol{u}_t$, $\boldsymbol{u}_t^1$ and of the contaminated shocks among the simulations, the average estimates of σ and σ_e^2 and the MSE1 and MSE2 measures computed for $Nsim = 100$ simulations for $\sigma_e^2 = 0.10, 0.25, 0.50$, $T = 100, 200, 2000$ and $N = 5, 10$. The SNR of $\boldsymbol{\alpha}_t$ and the SNR of the theoretical signal is also reported. Values in parentheses are the standard deviations.

	$\sigma_e^2 = 0.10$; $SNR_{\boldsymbol{\alpha}} = 10$; $SNR_{sys} = 11$											
N	5						10					
T	100		200		2,000		100		200		2,000	
$SNR_{\boldsymbol{u}}$	7.110	(.521)	7.039	(.349)	6.990	(.085)	4.065	(.215)	4.018	(.137)	4.000	(.039)
$SNR_{\boldsymbol{u}^1}$	11.183	(.920)	11.006	(.614)	10.991	(.162)	11.189	(.957)	10.950	(.633)	10.969	(.196)
$SNR_{\boldsymbol{\beta}}$	11.184	(.923)	11.011	(.628)	10.990	(.161)	11.180	(.917)	10.952	(.631)	10.969	(.195)
σ	1.000	(.018)	1.001	(.014)	1.000	(.004)	1.004	(.017)	1.003	(.013)	1.000	(.004)
σ_e^2	0.099	(.011)	0.099	(.007)	0.100	(.002)	0.098	(.005)	0.099	(.004)	0.100	(.001)
MSE1	0.801		0.271		0.145		0.809		0.263		0.147	
MSE2	0.009		0.009		0.001		0.010		0.006		0.003	
	$\sigma_e^2 = 0.25$; $SNR_{\boldsymbol{\alpha}} = 4$; $SNR_{sys} = 5$											
N	5						10					
T	100		200		2,000		100		200		2,000	
$SNR_{\boldsymbol{u}}$	3.441	(.227)	3.416	(.158)	3.395	(.039)	2.227	(.105)	2.208	(.067)	2.200	(.019)
$SNR_{\boldsymbol{u}^1}$	5.044	(.398)	4.993	(.273)	4.994	(.070)	5.008	(.417)	4.952	(.282)	4.985	(.089)
$SNR_{\boldsymbol{\beta}}$	5.037	(.403)	4.990	(.281)	4.994	(.070)	5.009	(.413)	4.952	(.285)	4.985	(.089)
σ	1.000	(.031)	1.003	(.024)	1.000	(.007)	1.011	(.028)	1.008	(.021)	1.001	(.006)
σ_e^2	0.247	(.026)	0.246	(.018)	0.250	(.006)	0.243	(.012)	0.246	(.009)	0.250	(.003)
MSE1	0.922		0.391		0.229		0.846		0.339		0.190	
MSE2	0.005		0.005		0.002		0.002		0.009		0.003	
	$\sigma_e^2 = 0.50$; $SNR_{\boldsymbol{\alpha}} = 2$; $SNR_{sys} = 3$											
N	5						10					
T	100		200		2,000		100		200		2,000	
$SNR_{\boldsymbol{u}}$	2.219	(.129)	2.208	(.093)	2.197	(.023)	1.614	(.064)	1.605	(.042)	1.600	(.012)
$SNR_{\boldsymbol{u}^1}$	2.995	(.223)	2.986	(.158)	2.995	(.040)	2.935	(.237)	2.945	(.162)	2.990	(.052)
$SNR_{\boldsymbol{\beta}}$	2.989	(.230)	2.982	(.161)	2.995	(.040)	2.931	(.234)	2.945	(.164)	2.989	(.052)
σ	1.004	(.051)	1.008	(.037)	1.000	(.012)	1.028	(.041)	1.019	(.031)	1.003	(.009)
σ_e^2	0.489	(.053)	0.489	(.036)	0.500	(.011)	0.483	(.024)	0.490	(.018)	0.499	(.006)
MSE1	1.115		0.783		0.429		0.991		0.643		0.283	
MSE2	0.004		0.003		1E-04		0.005		0.003		0.001	

pute MSE2 $= ||\widehat{A} - I_R||_F^2$, where $\widehat{A}$ is the average of $\widehat{A}_k$ among simulations. In Tables A.1 and A.2 are reported the MSE1 and MSE2 measures for the various parameter settings considered. Overall we may appreciate the consistency of the procedure when T becomes large regardless σ_e^2 and N.

We then graphically check the quality of the identification procedure obtained by

Table A.2 The average SNR of the observations $\boldsymbol{u}_t$, $\boldsymbol{u}_t^1$ and of the contaminated shocks among the simulations, the average estimates of σ and σ_e^2 and the MSE1 and MSE2 measures computed for $Nsim = 100$ simulations for $\sigma_e^2 = 0.10, 0.25, 0.50$, $T = 100, 200, 2000$ and $N = 20$. The SNR of $\boldsymbol{\alpha}_t$ and the SNR of the theoretical signal is also reported. Values in parentheses are the standard deviations.

	$\sigma_e^2 = 0.10$; $SNR_{\boldsymbol{\alpha}} = 10$; $SNR_{sys} = 11$					
N	20					
T	100		200		2000	
$SNR_{\boldsymbol{u}}$	2.529	(.076)	2.502	(.052)	2.500	(.019)
$SNR_{\boldsymbol{u}^1}$	10.948	(.908)	10.904	(.634)	10.950	(.175)
$SNR_{\boldsymbol{\beta}}$	10.958	(.928)	10.899	(.646)	10.948	(.175)
σ	1.014	(.020)	1.002	(.012)	1.001	(.005)
σ_e^2	0.097	(.003)	0.099	(.002)	0.100	(.001)
MSE1	0.700		0.305		0.176	
MSE2	0.003		0.002		4E-04	
	$\sigma_e^2 = 0.25$; $SNR_{\boldsymbol{\alpha}} = 4$; $SNR_{sys} = 5$					
N	20					
T	100		200		2000	
$SNR_{\boldsymbol{u}}$	1.614	(.041)	1.599	(.028)	1.600	(.010)
$SNR_{\boldsymbol{u}^1}$	4.833	(.372)	4.876	(.274)	4.972	(.080)
$SNR_{\boldsymbol{\beta}}$	4.829	(.377)	4.871	(.280)	4.971	(.079)
σ	1.035	(.033)	1.009	(.020)	1.002	(.007)
σ_e^2	0.242	(.008)	0.246	(.006)	0.250	(.002)
MSE1	0.852		0.418		0.166	
MSE2	0.003		2E-04		0.003	
	$\sigma_e^2 = 0.50$; $SNR_{\boldsymbol{\alpha}} = 2$; $SNR_{sys} = 3$					
N	20					
T	100		200		2000	
$SNR_{\boldsymbol{u}}$	1.309	(.027)	1.298	(.018)	1.300	(.006)
$SNR_{\boldsymbol{u}^1}$	2.771	(.198)	2.856	(.151)	2.978	(.047)
$SNR_{\boldsymbol{\beta}}$	2.763	(.200)	2.852	(.153)	2.977	(.047)
σ	1.077	(.047)	1.027	(.030)	1.004	(.011)
σ_e^2	0.479	(.015)	0.490	(.011)	0.499	(.004)
MSE1	1.062		0.674		0.304	
MSE2	0.012		0.019		0.003	

using the estimated mixing matrix $\widehat{Z}_k$ on the contaminated shocks $\boldsymbol{\beta}_t$. In particular, for each simulation path, we compute the identification on the contaminated shocks. we compare the average unmixed signals across simulations with the true sources $\boldsymbol{\alpha}_t$ using ID, see Figure A.1. On average, we observe that the separation of the sources is well performed also for the contaminated ones, since we can clearly distinguish the sources

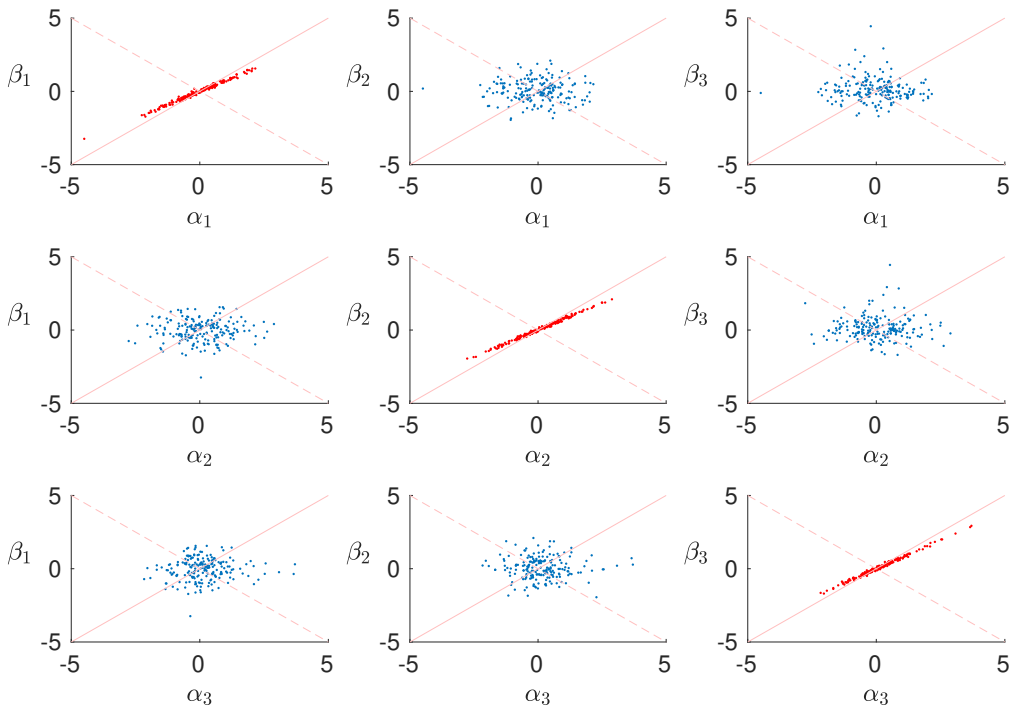


Fig. A.1 Scatter plots of shocks and contaminated shocks (average). This figure compares the shocks with the contaminated ones, which are obtained by the average of the contaminated shocks among the simulations. For each simulation path the identification scheme is applied to the obtained contaminated shocks. The scatter plot is expected to look like a straight line (along the bisector of the first-third quadrant) when the signal of the shock α_i is the same as that of β_j.

among the contaminated source signals[1].

By selecting one particular simulation, when $\sigma_e^2 = 0.5$, $T = 200$ and $N = 10$, we report the SNR of the sensors, virtual (collapsed) sensors and of the contaminated sources, see Figure A.2. Again, it is clear that the collapsing step can reduce the noise on the virtual (collapsed) sensors, whereas the source signal on the original sensors is weak. We report the plots of the shocks against the collapsed observations and the contaminated shocks, respectively. Figure A.3, panel (a), shows that the collapsed observations signal is basically a mixed signal of the original shocks, where all the information provided by the shocks is included. However, the collapsed observations are still an unknown linear combination of the original structural shocks so that the latter can not be identified yet. For instance, from the scatter plots on the right of panel (a) we observe that the signal of the collapsed observation u_3^1 is a combination of α_2 and $-\alpha_3$. So, we can identify each shock only after the identification process performed with the ICA on the collapsed observations, as it is exhibited in Figure A.3, panel (b). Indeed, from left to right,

[1]We also verify the goodness of the unmixing process for the unmixed source signals $\widehat{\boldsymbol{\alpha}}$ and we obtain similar results, even better due the absence of noise.

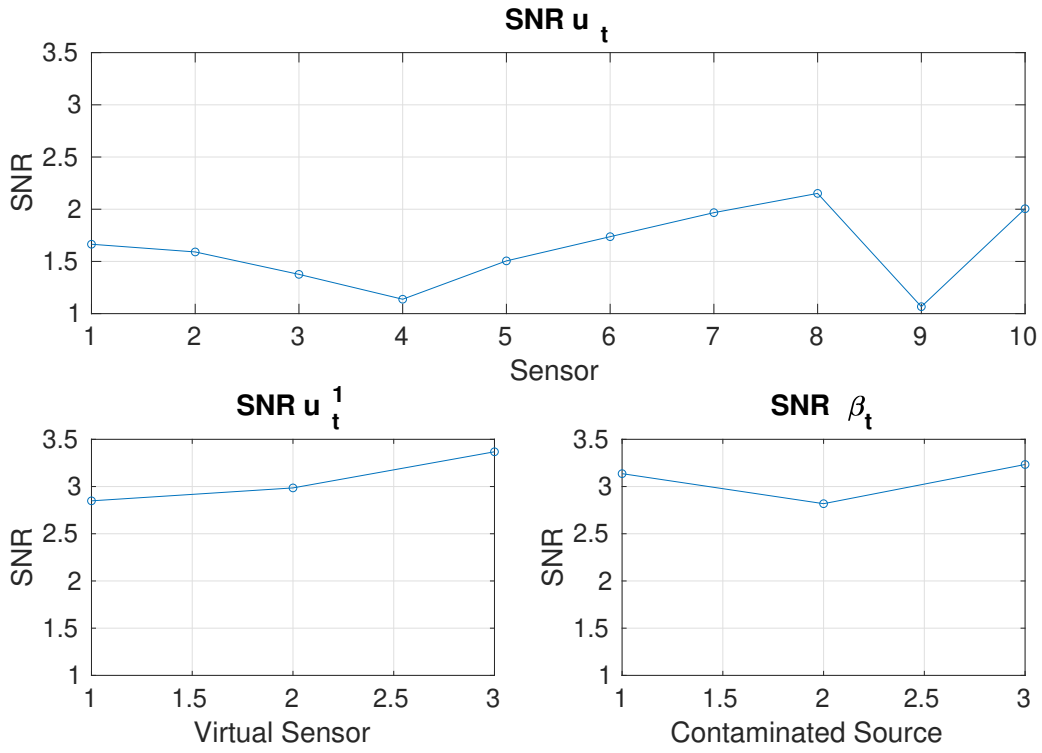


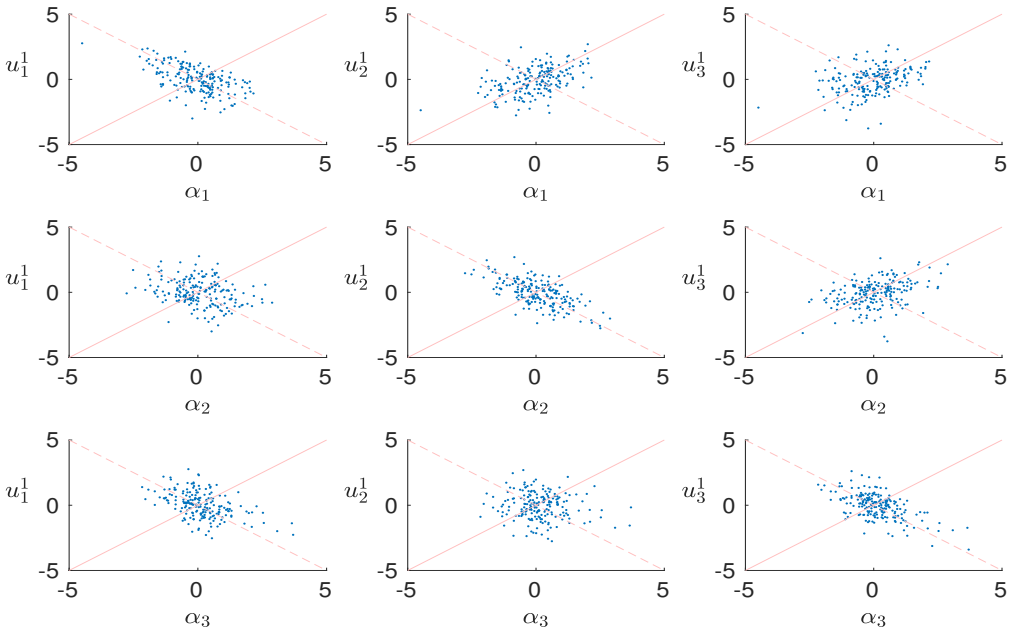
Fig. A.2 SNRs of one simulation for the original sensors, virtual sensors and for the unmixed contaminated sources. The collapsing procedure reduces the amplitude of the noise which is contained in the mixed sources signal.

we observe that signals in β_1, β_2 and β_3 almost perfectly coincide to $-\alpha_3$, α_1 and α_2, respectively.

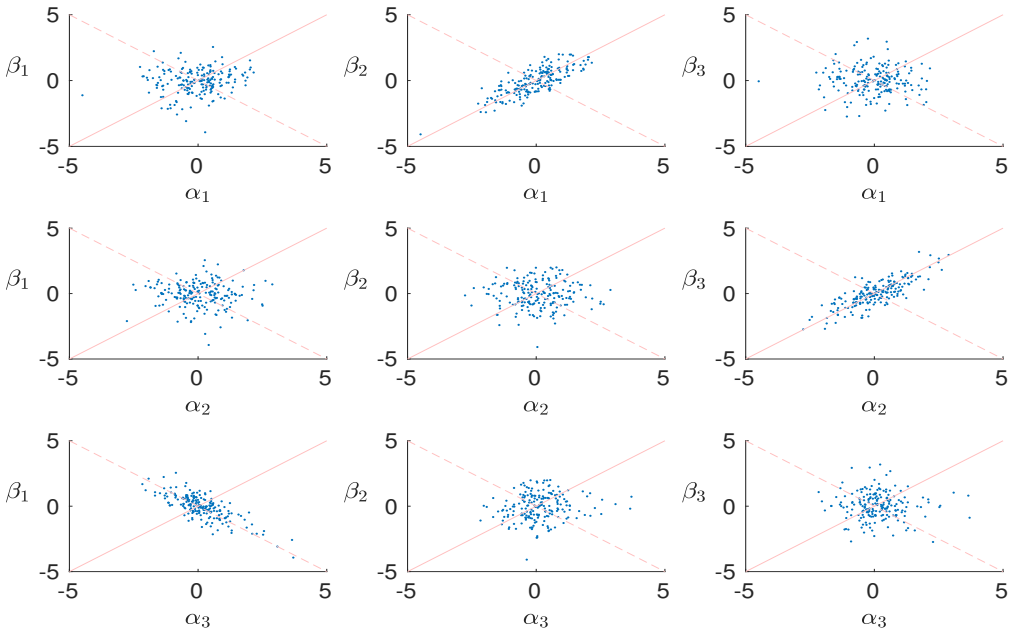
The ICA allows to separate the sources even when some noise remains in the mixed signal, since the ICA unmixing on the virtual sensors is sufficiently good to distinguish the sources making unnecessary a denoising step. Summarizing, the good accuracy of this identification process is achieved thanks to the preliminary collapsing step, which allows to considerably reduce the amplitude of the noise contained in the mixed shocks signal, and to the robustness of the ICA identification step to the remaining noise.

A.3 Estimated matrices of Section 2.3

We report the three mixing matrices estimated in the empirical analysis when all macroeconomic variables are considered. They are obtained from the short-run restrictions, PML-ICA and C-ICA approach. For the first two method we obtain a 4×4 matrix while for the C-ICA approach we estimate a rectangular 4×3 matrix. Table A.3 exhibits the structure of the obtained mixing matrices. For completeness, we report in Table A.4 the estimated mixing matrix $\widehat{C}'$ by the PML-ICA with the related statistic. We also perform the test of Gouriéroux et al. (2017) for testing whether the estimated $\widehat{C}'$ is in the equivalence class of the identity. The test rejects the null hypothesis with at any conventional confidence levels (p-value= 3.80e-05), i.e., $\widehat{C}' \notin \mathscr{P}(I_4)$. Furthermore, Table



(a)



(b)

Fig. A.3 Scatter plots of shocks and collapsed observations, panel (a), and scatter plots of shocks and contaminated shocks, panel (b), respectively. This figure compares the shocks with the collapsed observations and contaminated shocks for one particular simulation when $\sigma_e^2 = 0.5$, $T = 200$ and $N = 10$. The scatter plot is expected to look like a straight line (along the bisector of the first-third quadrant) when the signal of the shock α_i is the same as that of u_j^1 and β_j, respectively for panel (a) and (b).

Table A.3 Estimated mixing matrix $\widehat{Z}$ when short-run (Cholesky), PML-ICA and C-ICA are employed. We remark that for the short-run and PML-ICA the number of structural shocks is set equal to the number of the sensors, i.e., 4, while for the C-ICA procedure we set the number of the structural shocks equal to 3.

Short-Run	**PML-ICA**
$\begin{bmatrix} 0.8679 & 0 & 0 & 0 \\ -0.0211 & 0.2441 & 0 & 0 \\ 0.0307 & -0.3472 & 0.5764 & 0 \\ 0.1015 & -0.2218 & -0.0167 & 0.8685 \end{bmatrix}$	$\begin{bmatrix} 0.8026 & -0.0998 & 0.3088 & -0.0573 \\ 0.0219 & 0.2273 & -0.0272 & 0.0850 \\ -0.2224 & -0.2862 & 0.5711 & -0.0192 \\ 0.1303 & -0.5345 & -0.0670 & 0.7132 \end{bmatrix}$
C-ICA	
$\begin{bmatrix} -0.6981 & -0.3173 & -0.1972 \\ -0.0293 & -0.0099 & 0.1553 \\ 0.2624 & -0.1251 & -0.7205 \\ -0.2644 & 0.7145 & -0.2116 \end{bmatrix}$	

Table A.4 Estimated matrix $\widehat{C}'$ using the PML-ICA approach. Asymptotic standard deviations of the parameters are reported between parentheses and they are obtained using the asymptotic results of Gouriéroux et al. (2017).

$$\widehat{C}' = \begin{bmatrix} 0.9248 & -0.1150 & 0.3558 & -0.0661 \\ (0.0374) & (0.1227) & (0.0873) & (0.1289) \\ 0.1699 & 0.9212 & -0.0805 & 0.3426 \\ (0.1247) & (0.0591) & (0.1023) & (0.1494) \\ -0.3329 & 0.0645 & 0.9235 & 0.1766 \\ (0.0899) & (0.1072) & (0.0370) & (0.1060) \\ 0.0790 & -0.3655 & -0.1216 & 0.9198 \\ (0.1306) & (0.1477) & (0.1028) & (0.0604) \end{bmatrix}$$

A.5 reports the estimated collapsed mixing matrix $\widehat{C}'$ by the PML-ICA approach on the collapsed observations obtained by the collapsing step of the C-ICA approach. Even in this case the test of Gouriéroux et al. (2017) rejects the null hypothesis at any conventional confidence levels (p-value= $5e$-16). We observe that in contrast to the previous case, almost all the elements of $\widehat{C}'$ are statistically significantly different from zero. Table A.6 exhibits the mixing matrix obtained when the short-run (Cholesky) restrictions and PML-ICA are used for the scheme $[\pi, U_g, r]$. For this scheme, the matrix C' obtained with the PML-ICA is exactly the same estimated by Gouriéroux et al. (2017), since we have replicated the same exercise. We refer to their work for the statistics and related test.

Table A.5 Estimated matrix $\widehat{C}'$ using the PML-ICA approach on the virtual sensors provided by the collapsing step in the C-ICA procedure. Asymptotic standard deviations of the parameters are reported between parentheses and they are obtained using the asymptotic results of Gouriéroux et al. (2017).

$$\widehat{C}' = \begin{bmatrix} 0.6924 & -0.4856 & 0.5337 \\ (0.0635) & (0.0720) & 0.0768 \\ 0.5825 & 0.8126 & -0.0164 \\ (0.0657) & (0.0471) & (0.0879) \\ -0.4257 & 0.3222 & 0.8455 \\ (0.0832) & (0.0795) & (0.0483) \end{bmatrix}$$

Table A.6 Estimated mixing matrix $\widehat{Z}$ when short-run (Cholesky), PML-ICA are employed on the scheme $[\pi, U_g, r]$.

Short-Run	**PML-ICA**
$\begin{bmatrix} 0.8993 & 0 & 0 \\ -0.0388 & 0.2788 & 0 \\ 0.1209 & -0.3302 & 0.8958 \end{bmatrix}$	$\begin{bmatrix} 0.8593 & -0.2439 & -0.1042 \\ 0.0349 & 0.2779 & -0.0278 \\ 0.1580 & -0.2772 & 0.9079 \end{bmatrix}$

Table A.7 Variance Decomposition related to the 4-th structural shock. For each VAR model considered we report in columns the variance explained by the 4-th structural shock relative to each variables. The lag period after the shock is measured in quarters. Values in parentheses are the standard deviations computed from 100 bootstrap samples.

Lag	0		2		4		8		16	
$4SR$										
π	0	(0)	0.0166	(0.0148)	0.0217	(0.0186)	0.0293	(0.0292)	0.0275	(0.0427)
U_g	0.9936	(0.0170)	0.8223	(0.0570)	0.6839	(0.0910)	0.6101	(0.1202)	0.5499	(0.1339)
O_g	0.2703	(0.0567)	0.3074	(0.0778)	0.2982	(0.0914)	0.3147	(0.1178)	0.3122	(0.1248)
r	0.0523	(0.0447)	0.2002	(0.0801)	0.2388	(0.0884)	0.2342	(0.0994)	0.2254	(0.1073)
$4PML\text{-}ICA$										
π	0.0132	(0.0383)	0.0365	(0.0398)	0.0443	(0.0425)	0.0422	(0.0541)	0.0314	(0.0924)
U_g	0.8594	(0.0545)	0.6073	(0.0562)	0.4317	(0.0784)	0.3158	(0.1182)	0.2524	(0.1465)
O_g	0.1789	(0.0448)	0.1579	(0.0641)	0.1339	(0.0795)	0.1193	(0.1121)	0.1127	(0.1301)
r	0.3502	(0.1312)	0.4617	(0.1248)	0.4609	(0.1092)	0.4173	(0.0991)	0.3666	(0.0864)

Finally, Table A.7 reports the variance decomposition related to the 4-th structural shock for the 4-VAR model using short-run restrictions and PML-ICA approach.

CHAPTER 3

Uncertainty in Firm Valuation

In valuation analysis, the degree of uncertainty associated with the value of a company is as important as the value itself. We propose an original and robust methodology for company market valuation which replaces the traditional point estimate of the conventional Discounted Cash Flow model with a probability distribution of fair values that conveys information about both the expected value of the company and its intrinsic uncertainty. The methodology hinges on two main ingredients: an econometric model for the company revenues and a set of firm-specific balance sheet relations that are estimated using historical data. In following chapters, Chapters 4 and 5, we explore the effectiveness and scope of our methodology through a series of statistical exercises on publicly traded U.S. companies. In this Chapter, we describe the proposed method by presenting the main assumptions and its main building blocks.

3.1 The Stochastic Discounted Cash Flow model

The Stochastic Discounted Cash Flow model generalizes the two-stage UFCF DCF model presented in Chapter 1. In what follows, we assume that all the random quantities are defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathcal{P})$. In addition, we denote with small (resp. capital) letters all the deterministic (resp. random) quantities. In the standard DCF method, the (present) value v_0 of a firm is defined by the sum of the expected discounted cash flows, see Eq. (1.2), and it is a single point estimate. The uncertainty associated to the future firms' performance and to the evaluation process itself is summarized by the expected value of future cash flows. In SDCF, the single point estimate v_0 is replaced by the realization of a random variable V_0 such that $\mathbb{E}[V_0] = v_0$. In particular, in this framework, Eq. (1.2) can be rewritten $\forall \omega \in \Omega$ as,

$$V_0(\omega) = \sum_{t=1}^{T} \frac{CF_t(\omega)}{(1+k)^t} + \frac{CF_T(\omega)(1+g)}{(1+k)^T(k-g)} \tag{3.1}$$

and the corresponding equity fair value, V_0^{Eq} is obtained by replacing v_0 by V_0 in Eq. (1.4). The distribution of the equity fair value divided by the number of company's outstanding shares is called *fair value distribution*, which can be compared with the corresponding

stock price[1].

The previous equation implies that the main building block in the definition of V_0 is a reliable description of the future cash flow generating process CF_t. However, the cash flow is on its own a reconstructed accounting variable, which depends on a set of other more fundamental variables. We recall that our measure of cash flows is constituted by the UFCF defined in Eq. (1.3) and in the following, for the sake of simplicity, we will denote it with CF.

The interaction between the accounting variables and the cash flow is manifold and so, in order to obtain a manageable model, we employ the following two-steps procedure. First, all the relevant accounting variables are expressed as *margins* with respect to the revenues. These margins are estimated in a firm-specific regression framework. Second, an econometric model for the revenues' dynamic is proposed. This model is calibrated at the single firm level and is used in a Monte Carlo exercise in order to obtain a statistical description of the future revenues (and therefore of the future cash flow). The rationale behind this approach is that on the one hand, the relation between the relevant accounting variables and revenues constitutes a structural relationship that is likely to remain stable across time. On the other hand, the revenues are an important idiosyncratic component at the single firm level, besides being an important driver of the growth of a firm. Indeed, this is exactly what we observe from the study of Section 1.2, when we analyze the causal properties using the SVAR approach, which is also consistent with the results of Coad and Rao (2010), Moneta et al. (2013).

We now first construct a reliable relation between the cash flow and the revenues, then we focus on the econometric model for the latter quantity. Discounting rates are discussed at the end of this chapter.

3.1.1 Cash flows and revenues

From Eq. (1.3), we define for each time t, $\overline{CF}_t$ as the sum between $NOPAT_t$ and $D\&A_t$ less $CAPEX_t$, i.e., $\overline{CF}_t = NOPAT_t + D\&A_t - CAPEX_t$. Therefore, $CF_t = \overline{CF}_t - \Delta WC_t$ and we express each accounting variable in function of revenues. Nonetheless, $NOPAT_t$, $D\&A_t$, $CAPEX_t$ and WC_t are customarily expressed as marginal quantities upon the revenues REV_t, i.e.,

$$\overline{CF}_t = \alpha \cdot REV_t, \quad WC_t = \beta \cdot REV_t$$

so that

$$CF_t := (\alpha - \beta) \cdot REV_t + \beta \cdot REV_{t-1}.$$

[1] Since, this distribution is estimated by a Monte Carlo procedure, in the following we will consider the *empirical fair value distribution* instead of the true one. However, we will not make this distinction for the sake of simplicity.

We model $\overline{CF}_t$ and WC_t with two separate margins, since a distinguishing difference between $\overline{CF}_t$ and ΔWC_t is the frequency at which they are updated. Precisely, the former is updated every quarter by using trailing twelve months data, while the latter refers to annual variables. Then, we estimate α and β in the previous equations through two different steps. First, we estimate α with the following regression of $\overline{CF}_t$ on REV_t:

$$\begin{cases} \overline{CF}_t & = \alpha\, REV_t + u_t \\ u_t & = \varepsilon_t - \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}, \quad \varepsilon_t \overset{d}{\sim} \mathcal{N}(0, \sigma^2), \end{cases} \tag{3.2}$$

where $\mathcal{N}(\mu, \sigma^2)$ denotes a Gaussian random variable with mean μ and variance σ^2. The number of lags q is fixed by using standard information criteria (e.g., the Akaike Information Criterion (AIC)). In the empirical analysis, see Chapter 4 and 5, regression in Eq. (3.2) is first estimated over the time period FQ4 1992-FQ1 2009 and subsequently over the remaining periods by shifting the initial estimation window of one fiscal quarter. Second, following our data provider Thomson Reuters Eikon, Datastream database, we set the operating margin of WC_t on revenues equal to the annual sample average of the historical working capital margin rate computed initially over the period 2006–2009, i.e., over the last three years and subsequently by shifting the initial window of one year.

Hereafter, for sake of presentation, the description of the procedure is related to a specific period only. Letting $\hat{\alpha}$ and $\hat{\beta}$ be these estimates, then our measure of cash flow is given by

$$CF_t = (\hat{\alpha} - \hat{\beta})\, REV_t + \hat{\beta}\, REV_{t-1}, \tag{3.3}$$

i.e., the cash flows are characterized in terms of revenues' dynamic. The coefficient of determination of the previous regressions is on average greater than 0.9. More details on model checking assessing the reliability of the previous relations is reported in Section 3.3.1. However, analysts might be used their best model to forecast future cash flows, and we remark that the proposed framework is invariant from this choice. This point it will be discussed in Chapter 5, where we argument that what is really important to issue sell-side recommendations is to have a good tool to evaluate at the same level the different companies using a sort of relative comparison in the selected stocks universe.

3.1.2 The revenues' process

The revenues dynamics of each company is described by comparing three alternative econometric model. It is important to stress that we are not searching for a revenues' common-structure to be identified in a cross-sectional fashion. Rather, we estimate the three models on each company separately and then, independently for each firm, we select the best model. Indeed, we are interested in a firm-specific model for the revenues. This kind of way of thinking is standard in the DCF valuation model, since it is implicitly

assumed that the value of a firm is fully characterized by their fundamentals. However, in the proposed SDCF model is highlighted how it is actually important to have a good model of the driver process. Thus, the best model, with the best out-of sample performance, should be employed. However, this research can be very challenging and it is beyond the scope of this thesis. Therefore, we focus on simple models which outperform, for instance, the simple AR(1) model.

Because trailing twelve-months data are used, we neglect seasonal characteristics. The three alternative models are the following. Let $y_t = \log(REV_t)$ be the logarithm of the revenues,

a) Model 1 or stationary model:

$$\left(1 - \sum_{k=1}^{p} a_k L^k\right)(1-L)y_t = \varepsilon_t, \quad \varepsilon_t \overset{d}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2),$$

where L is the usual lag operator, i.e. $Ly_t = y_{t-1}$.

b) Model 2 or local level model:

$$\begin{aligned} \text{(Observation equation)} \quad & y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \overset{d}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2) \\ \text{(Local level)} \quad & \mu_{t+1} = \mu_t + \eta_t, \quad \eta_t \overset{d}{\sim} \mathcal{N}(0, \sigma_\eta^2), \end{aligned}$$

where the shocks ε_t and η_t are independent.

c) Model 3 or local linear trend model:

$$\begin{aligned} \text{(Observation equation)} \quad & y_t = \mu_t + \varepsilon_t, && \varepsilon_t \overset{d}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2) \\ \text{(Local trend)} \quad & \mu_{t+1} = \mu_t + \nu_t + \eta_t, && \eta_t \overset{d}{\sim} \mathcal{N}(0, \sigma_\eta^2) \\ \text{(Time-varying slope)} \quad & \nu_{t+1} = \nu_t + \zeta_t, && \zeta_t \overset{d}{\sim} \mathcal{N}(0, \sigma_\zeta^2), \end{aligned}$$

where the three innovations ε_t, η_t and ζ_t are independent.

Model 1 assumes that the first difference of log-revenues are stationary and described by an AR(p) model, with the lag p selected (again) according to AIC. Model 2 and Model 3 assume that the log-revenues y_t depend upon non observable state variables, the local trend μ_t and the time varying slope ν_t, and they represent two alternative models when the differenced time-series of revenues, Δy_t, is not stationary. These models are estimated in their state-space form using the Kalman filter, which is used to estimate the parameters of the models and the time series of the unobservable state variables (see, e.g., Harvey (1990) and Durbin and Koopman (2012) for further details). The model selection procedure is made of two steps. First, we check if the time series of the log-revenues' increments is stationary. If it is the case, we select Model 1. If it is not the case, we

estimate both Model 2 and Model 3. Then, since Model 2 is nested in Model 3, we use the likelihood ratio test to select the best between the two. Once one model among the three has been selected, we compute a forecast of future log-revenues, and therefore of future cash-flow, by employing a bootstrap procedure. Section 3.3.2 reports analysis on the goodness of fit of the three models together with a performance comparison against the AR(1) model.

3.1.3 The discounting and perpetual growth rate

Let us clarify the non trivial issue concerning the choice of the discount rate of Eq. (3.1) when CF_t is a stochastic process. In the conventional DCF model the expected future cash flows are discounted back using the cost of capital rate, which incorporates the intrinsic uncertainty of the (future) cash flows. Therefore, should the model for the CF_t account in full for the intrinsic uncertainty of the latter, the discount rate would be the riskless rate (see, e.g., Casey, 2001). However, this is true only when the specified cash flows model has been chosen in such a way that all the intrinsic uncertainty is considered in all its components, i.e., the cash flows are risk-neutral. In other words, one have to discount all the uncertainty which is not explained by the CF_t process. In the particular case of the presented SDCF model, the cash flows of a specific firm are driven only by the idiosyncratic shocks of the revenues process[2], which is related only to the specific dynamic involved by the selected firm, without accounting, for instance, any possible co-integration with other firms. In particular, since the CF_t model is calibrated at level of the single firm with univariate time series models, we do not to consider any other form of risk, such as the systematic one, and we have not eliminate all the risk related to the company's business life. Then, we continue to use the same discount rate employed in the conventional DCF in Eq. (1.2) also in Eq. (3.1)[3]. Moreover, from mathematical point of view this implies a nice property which relates the SDCF to the pointwise DCF, i.e., the $\mathbb{E}[V_0]$ is exactly the value obtained by the pointwise DCF v_0, which means that with the SDCF the analyst is able to obtain a confidence interval of v_0.

In conclusion, if one selects the "true" cash flows model which incorporates all the uncertainty, then one have to use the riskless rate as discount rate. If the model does not incorporate all the uncertainty, then the analyst have to discount that uncertainty which is not included in such cash flow model.

[2]More precisely, the cash flows are driven by the revenues process which in turn is driven by idiosyncratic shocks.

[3] We also observe that this choice agrees with that made in Ali et al. (2010), Razgaitis (2009) chapter 8; French and Gabrielli (2005) Dayananda et al. (2002) where the use of Monte Carlo simulation in DCF is discussed.

3.1.4 The stochastic valuation framework

The general framework which we proposed for firm valuations using a SDCF approach is characterized by the following three step.

a) Determine the most appropriate model for cash flows.

b) Compute each future trajectory of $CF_t(\omega)$, $\omega \in \Omega$.

c) Discount back these trajectories with the appropriate discount rates and compute the aggregate fair value distribution.

In particular, in our specific SDCF methodology the key steps a) and b) are substituted by the following points:

i) Estimate the slope $\hat{\alpha}$ through the time-series regression (3.2) and the operating margin $\hat{\beta}$ of WC_t upon the revenues.

ii) Determine the most reliable model for the log-revenues, selecting one among Model 1, Model 2 and Model 3.

iii) Compute each future trajectory of $CF_t(\omega)$, $\omega \in \Omega$ in Eq. (3.3) by replacing the trajectory for the revenues process, with the bootstrapped value $\mathrm{REV}_t(\omega)$ determined by either Model 1, Model 2 or Model 3.

Figure 3.1 shows two examples of the logarithm of the fair value distribution for Booking Holdings Inc. (ticker BKNG) and McCormick & Company (ticker MKC) computed at different dates; the red dotted lines indicate the market price at the evaluation date.

3.2 Data and sample selection criteria

While the SDCF model relies upon general considerations, many details of its implementation and validation depend on specific company level data. In this Section we review the different data sources we use to develop and test our methodology. All the data in the subsequent description are taken at a quarterly frequency.

The required equity prices, along with the corresponding fundamental data are collected from Thomson Reuters Eikon, Datastream database. Our initial sample comprises all non-financial firms included in the S&P 500 over the entire period January 2009-December 2017, without missing data, and with sufficient observations of the revenues before January 2009 to estimate our model[4] In this sample of 182 firms, we discard 32

[4]We exclude financial companies as they are subject to industry-specific regulations and present a peculiar capital structure which makes very difficult, if not meaningless, the estimation of cash flows. In general, missing data for company fundamentals (e.g., EBITDA and CAPEX) complicates the reconstruction of the past free cash flow and/or the computation of the present value.

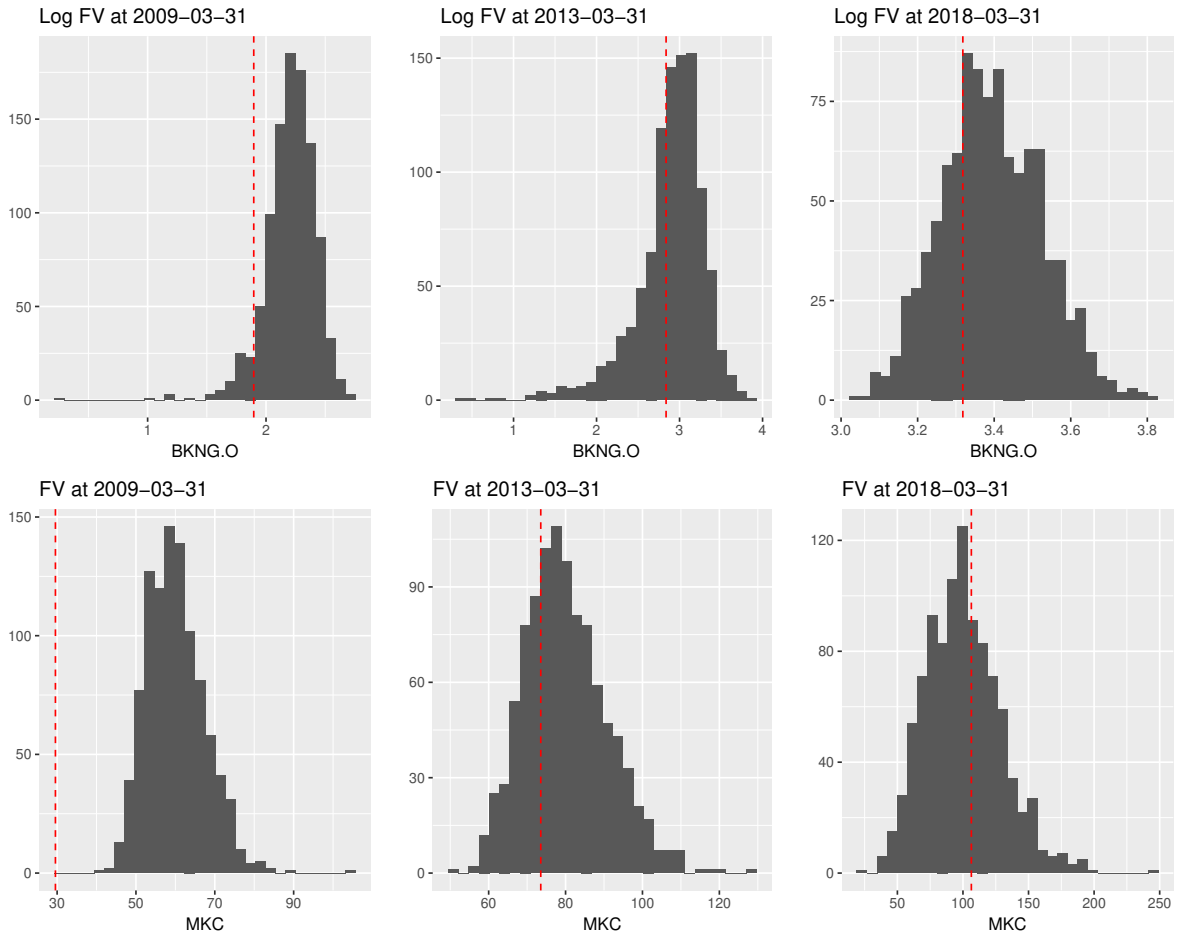


Fig. 3.1 Examples of the *fair value distribution* (labelled FV) for Booking Holdings Inc. (ticker BKNG) and McCormick & Company (ticker MKC) computed at different dates. The red dotted lines indicate the market price at the evaluation date. For Booking Holdings Inc. the logarithm of the fair value distribution is displayed.

firms for which the coefficient of determination (R^2) of the regression of $\overline{CF}_t$ on REV_t results to be less than 10% and another 10 firms for which we observe that for specific quarters, immediately after the financial crisis of 2008-2009, their estimated fair value distribution is negative.[5] All together, these modifications leave us with a sample of 140 firms.[6]

As regards the heterogeneity, among the 140 firms we have[7]: 17 firms in both the Oil & Gas (ICB 1) and the Basic Material (ICB 1000) sector, 44 Industrial firms (ICB 2000), 22 Consumer Good firms (ICB 3000), 19 Healthcare firms (ICB 4000), 12 firms in the Consumer Service sector (ICB 5000), 3 firms in the Telecommunication sector (ICB 6000), 7 Utilities firms (ICB 7000) and 16 Technology Firms (ICB 9000). Table 3.1

[5]This is attributed to the bad performance during the crisis which affect the correct estimation of the true latent driver process on the next quarter.

[6]Since we are interested in the comparison of companies within our selected sample of stocks, the survivorship bias which may be induced by the sample selection criteria does not compromise our analyses.

[7]We use the Industry Classification Benchmark (ICB) as industry classification taxonomy.

Table 3.1 Percentage of the stocks of our universe relatively to the S&P500 for each ICB sector.

ICB sector	(%)	ICB sector	(%)
1 and 1000	36.17	5000	16.90
2000	49.44	6000	100.00
3000	37.93	7000	25.00
4000	36.54	9000	30.19

Table 3.2 Proportion of stocks in each ICB sector relative to both our universe (first and second column) and the S&P500 (third and fourth column).

Our Universe				S&P 500			
ICB sector	(%)	ICB sector	(%)	ICB sector	(%)	ICB sector	(%)
1 and 1000	12.14	5000	8.57	1 and 1000	11.72	5000	17.71
2000	31.43	6000	2.14	2000	22.19	6000	0.75
3000	15.72	7000	5.00	3000	14.46	7000	6.98
4000	13.57	9000	11.43	4000	12.97	9000	13.22

Table 3.3 Summary statistics for the 140 firms in our sample, grouped by ICB code. Two "snapshots" are reported: one referring to January 2009, and one to December 2017 (between brackets), which correspond to the beginning and the end date of the period under investigation. For each sector firm is reported the average and the median values of market capitalization, revenues and net income.

	Oil & Gas and Basic Materials	Industrial	Consumer Goods	Healthcare	Consumer Service	Telecom. and Utilities	Technology	All
ICB codes	1 and 1000	2000	3000	4000	5000	6000 and 7000	9000	
Number of firms	17	44	22	19	12	10	16	140
Market cap. ($B)								
Mean	47.9 (62.7)	13.5 (35.9)	26.0 (53.2)	34.8 (87.0)	17.3 (104.1)	34.1 (59.2)	51.8 (165.8)	28.7 (71.1)
Median	11.1 (27.4)	5.9 (22.4)	6.0 (19.2)	16.4 (54.4)	7.0 (25.8)	5.6 (18.7)	36.4 (70.2)	7.7 (25.6)
Revenues ($B)								
Mean	69.8 (38.6)	15.5 (16.2)	16.5 (18.6)	17.1 (28.5)	21.3 (50.3)	28.9 (39.1)	28.3 (41.8)	25.4 (28.4)
Median	17.8 (15.9)	6.3 (9.8)	7.2 (13.3)	6.7 (13.0)	16.3 (19.1)	9.1 (13.3)	9.6 (18.4)	7.8 (12.4)
Net Income ($B)								
Mean	3.0 (1.8)	1.3 (1.2)	1.6 (2.5)	2.3 (3.1)	1.2 (2.9)	1.9 (2.6)	4.3 (8.1)	2.1 (2.8)
Median	0.4 (0.9)	0.4 (0.7)	0.5 (1.0)	1.1 (1.3)	0.6 (1.2)	0.5 (1.0)	2.1 (2.5)	0.5 (1.0)

reports for each ICB sector the percentage of stocks of our universe relative to the number of firms of the same sector in the S&P 500 index. To check for possible sample distortion introduced by our selection criteria, Table 3.2 displays the percentage of stocks in each ICB sector relative to both our universe (first two columns) and the S&P 500 (second two columns). Together, Tables 3.1 and 3.2 show that the final sample presents a great heterogeneity in terms of industry sectors and, importantly, reflects the composition of the index.

Table 3.3 presents summary statistics for our sample. Two "snapshots" are considered: one referring to the beginning of the period, i.e. January 2009, and one to the end, i.e. December 2017. Data relative to December 2017 are reported between brackets. The table shows average and median values across the firms within the same sector of market capitalization, revenues and net income.

The data employed in the analysis of the returns of different portfolios within standard multi-factor models (e.g., the market excess return and the momentum factor) are taken from the Kenneth R. French Data Library[8]. In addition, since we set the terminal year T to 5, the perpetual growth rate g is determined by the 5-year T-bond rate obtained from the FRED (Federal Reserve Bank in St. Louis) database. Finally, the corporate tax rate for United States firms is provided by KPMG.

Remark 3.1. Let us clarify a technical issue regarding the discounting rate of the terminal value. Following an industry standard, e.g., the Thomson Reuters Eikon methodology, the discount rate for the cash flow terminal value is computed by considering the fixed corporate tax rate instead of the individual tax rate, albeit the difference is minimal for all firms and all years considered. For the sake of completeness, if the discounting rate related to the terminal value is denoted by k_{TV}, then $TV = \frac{CF_T(1+g)}{(k_{TV}-g)}$ and it is discounted by $(1 + k_{TV})^T$. For the reader convenience, Table 3.4 reports summary statistics of the WACC for each quarter, as well as of the perpetual growth rate.

3.3 Model validation

The SDCF model is based on the UFCF DCF model employed by the data provider Thomson Reuters Eikon, Datastream, which is considered as a benchmark. However, we test the reliability of the choices made in the model trhough statistical and econometric analysis.

3.3.1 Statistical validation of cash flows model

The coefficient of determination of regression Eq. (3.2) is on average greater than 0.90, which assesses the goodness of fit of the linear regression of $\overline{CF}_t$ on REV_t. As regards the model misspecification, we check the normality and the autocorrelation of the standardized residuals. The Kolmogorov-Smirnov test fails to reject the assumption of normality of the standardized residuals for about the 56% of the firms in our universe at 0.01 level. Also, there is no evidence of serial correlations for about 72% of the firms using the Ljung-Box statistic.

[8] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Table 3.4 Summary statistics of the WACC, k, for each quarter, as well as of the perpetual growth rate, g. Standard deviation of k are reported in parenthesis.

	2009-03-30	2009-06-30	2009-09-30	2009-12-30	2010-03-30
Average $WACC$	7.53	7.85	7.89	8.01	6.17
	(1.80)	(1.82)	(1.84)	(1.83)	(1.28)
g	2.23	2.47	2.3	2.42	2.25
	2010-06-30	2010-09-30	2010-12-30	2011-03-30	2011-06-30
Average $WACC$	5.53	5.52	6.08	6.67	6.07
	(1.19)	(1.19)	(1.23)	(1.51)	(1.47)
g	1.55	1.49	2.12	1.86	1.15
	2011-09-30	2011-12-30	2012-03-30	2012-06-30	2012-09-30
Average $WACC$	5.81	5.81	6.44	6.29	6.34
	(1.44)	(1.46)	(1.75)	(1.70)	(1.74)
g	0.95	0.9	0.79	0.67	0.69
	2012-12-30	2013-03-30	2013-06-30	2013-09-30	2013-12-30
Average $WACC$	6.48	6.31	6.81	6.78	6.81
	(1.81)	(1.79)	(1.81)	(1.80)	(1.66)
g	0.83	0.92	1.51	1.44	1.6
	2014-03-30	2014-06-30	2014-09-30	2014-12-30	2015-03-30
Average $WACC$	6.16	6.2	6.09	5.97	6.68
	(1.44)	(1.36)	(1.32)	(1.34)	(1.52)
g	1.66	1.7	1.6	1.45	1.52
	2015-06-30	2015-09-30	2015-12-30	2016-03-30	2016-06-30
Average $WACC$	6.72	6.57	6.44	6.51	6.45
	(1.55)	(1.40)	(1.44)	(1.51)	(1.53)
g	1.55	1.59	1.37	1.24	1.13
	2016-09-30	2016-12-30	2017-03-30	2017-06-30	2017-09-30
Average $WACC$	6.82	7.1	6.62	6.65	6.88
	(1.43)	(1.47)	(1.34)	(1.36)	(1.40)
g	1.61	1.94	1.81	1.82	2.07
	2017-12-31				
Average $WACC$	7.26				
	(1.42)				
g	2.54				

3.3.2 Goodness of fit of revenues models

We report an analysis on the goodness of fit of the models proposed for the log-revenues (Model 1, Model 2 and Model 3), see Section 3.1.2, and a comparison against the simple autoregressive model of order one AR(1). In order to do so, we proceed as follows, where we also quickly recall the model selection:

i) If the time series of the log-revenues' increments is stationary we use Model 1 and the performance of ARIMA-based models is assessed by looking at the usual coefficient of the determination.

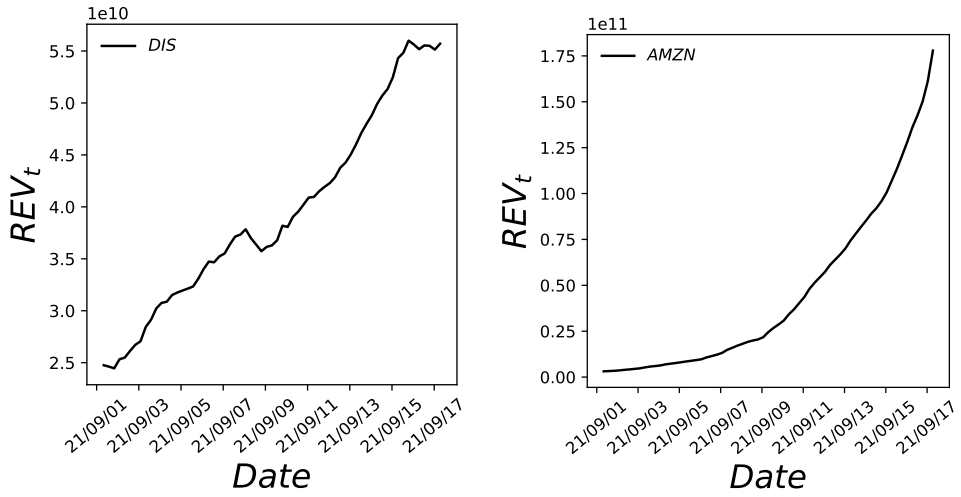


Fig. 3.2 Examples of the dynamic of the Revenues well fitted by Model 1 (*Left Panel*) and Model 3 (*Right Panel*). Model 1: Disney (ticker DIS). Model 3: Amazon (ticker AMZN).

ii) On the contrary, in case of non-stationarity, i.e., either Model 2 or Model 3 is selected, we proceed in the following way. Let $\bar{y}$ be the sample average of the log-revenues computed over a predetermined time period, and $\mu_{t|pred} = \mathbb{E}[\mu_t|\mathcal{F}_{t-1}]$ be the Kalman filter prediction of the state variable μ at time t. If T is the length of the considered time-series and K the number of explanatory variables, then we define the *adjusted* coefficient of determination R^2_{adj} as:

$$R^2_{adj} := 1 - (1 - R^2)\frac{T-1}{T-K-1}, \quad \text{where} \quad R^2 := 1 - \frac{SS_{res}}{SS_{tot}},$$

and $SS_{res} = \sum_{t=1}^{T}(y_t - \mu_{t|pred})^2$ and $SS_{tot} = \sum_{t=1}^{T}(y_t - \bar{y})^2$.

The following results are obtained by the average on the whole sample considered (FQ4 2008- FQ4 2017). For the 71% of the firms in the sample Model 1 is the model that present the best goodness of fit. Of the remaining, only the 1.5% of firms have a log-revenues process that is well fitted by Model 2. Regardless the model selection we found that the coefficient of determination is greater than 0.90 on average among the sample period and the firm cross-section.

The plots in Figure 3.2 report the typical dynamics of the revenues process in Model 1 (*left panel*) and Model 2-3 (*right panel*). Examples of firms for which the latter models have been selected are Amazon (ticker AMZN) and Booking (ticker BKNG).

Models 2 and 3 assume that shocks are *i.i.d.* with a zero-mean Gaussian distribution. Following Durbin and Koopman (2012) we check for a possible misspecification of these models by looking at the time series of the standardized one-step ahead forecast errors $\hat{\eta}$ and we refer to Durbin and Koopman (2012) for further details on this definition.

In case of correct specification of the models, the latter should be *i.i.d.* and normally distributed. In detail, the 74% of the firms in the sample exhibit $\hat{\eta}_t$ with a distribution that is not statistically different from the Gaussian at the 0.01 confidence level according to the Kolmogorov-Smirnov test. Also, the $\hat{\eta}_t$ of the 76% (resp. 78%) of the firms does not exhibit significant serial autocorrelation at the 0.01 confidence level according to the Ljung-Box test performed at lag 1 (resp. 10). We argue that the proposed framework represents an optimal trade-off between a model with a perfect statistical fit of the data and a model that have clear advantages in terms of practical implementability.

Finally, we present results on the predictive ability of Model 1, Model 2 and Model 3 for our sample firms by comparing it with that of a simple autoregressive model of order one AR(1), which is described by the following dynamic: $y_t = \alpha y_{t-1} + \varepsilon_t$ where $|\alpha| \leq 1$ and $\varepsilon_t \overset{d}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2)$. We start by estimating all models over the period December 1992 – March 2009 and by making predictions for the next 20 quarters (i.e. for the next 5 years). We then add the most recent observation, we shift the estimation period by one quarter and we make predictions for the next 20 quarters, and so on and so forth till the first quarter of 2013.[9] We compute the following (relative) error metric for each firm j and each estimation period e:

$$\mathrm{RMSQ}_j^e := \frac{\mathrm{MSE}_{j,our}}{\mathrm{MSE}_{j,\mathrm{AR}(1)}},$$

where MSE stands for Mean Square Error and *our* is either Model 1, Model 2 or Model 3, depending on the selected model. The mean square error is defined in the natural way as $\mathrm{MSE} := \frac{1}{H}\sum_{h=1}^{H}(y_{t+h} - \hat{y}_{t+h})^2$, where $H = 20$ and y and $\hat{y}$ denotes the actual and the predicted value, respectively. For conciseness, for each estimation period e we compute the average across-firms of the RMSQ_j^e. In addition, we test whether our models perform better than the AR(1): a negative average of $\log(\mathrm{RMSQ}_j)$ means that our models have a MSE smaller than the that of the AR(1). Table 3.5 reports the average of the logarithm of RMSQ_j among the quarters. We observe that the average $\log(\mathrm{RMSQ}_j^e)$ is always negative and statistically significant for all quarters, with an exception of the last quarters of 2009 and the first of 2010, thus confirming that a more complex dynamics than the simple AR(1) one is necessary. Also, for each quarter, we compute the percentage of firms for which $\mathrm{RMSQ}_j < 1$ and the results are reported in Table 3.6. On average, for 65% of companies our models perform better than the AR(1).

We have presented a generalization of the traditional DCF model of firm valuation, in which the point estimate is replaced with an estimated probability distribution of fair values. Therefore, we can derive both an estimate of the fair value of a company and a measure of the degree of uncertainty associated with it. The statistical analysis shows the

[9] We remind that the data sample for the revenues end in the first quarter of 2018.

Table 3.5 Average $\log(\text{RMSQ}_j)$ computed across firms for all the quarters. Values that result to be significant at 10%, 5%, 1% and 0.1% level are marked with '$\circ$', '$*$', '$**$' and '$***$' respectively.

2009-03-31	2009-06-30	2009-09-30	2009-12-31	2010-03-31	2010-06-30
-0.61 ***	-0.33 *	0.31 *	0.43 **	0.14	-0.72 ***
2010-09-30	2010-12-31	2011-03-31	2011-06-30	2011-09-30	2011-12-31
-0.81 ***	-0.89 ***	-0.84 ***	-0.82 ***	-0.87 ***	-0.78 ***
2012-03-31	2012-06-30	2012-09-30	2012-12-31	2013-03-31	
-0.59 ***	-0.54 **	-0.66 ***	-0.59 ***	-0.68 ***	

Table 3.6 Percentage of firms j such that $\text{RMSQ}_j < 1$ among the all quarters. On average 65% companies have a RMSQ which is smaller than 1.

2009-03-31	2009-06-30	2009-09-30	2009-12-31	2010-03-31	2010-06-30
65 %	58.57 %	45.71 %	44.29 %	48.57%	70.00 %
2010-09-30	2010-12-31	2011-03-31	2011-06-30	2011-09-30	2011-12-31
74.29 %	72.14 %	75.71 %	72.14 %	74.29 %	68.57 %
2012-03-31	2012-06-30	2012-09-30	2012-12-31	2013-03-31	
63.57 %	65.71 %	67.86 %	64.29 %	68.57%	

performance of the considered econometric approaches and how they are well-specified among the selected universe of stocks. However, in the following chapters we will investigate the effectiveness of our methodology and show the importance of a distributional approach to valuation through different exercises. We will explore the misvaluation effect in Chapter 4 and in Chapter 5 we will propose two recommendation systems based on the comparison of observed market prices with the fair value distributions obtained through the introduced SDCF method.

CHAPTER 4

A Cross-Sectional Misvaluation Measure and the Valuation Factor

We now introduce a measure of misvaluation based on the fair value distribution information. Precisely, at firm level, we show that the fair value distribution derived with SDCF constitutes a reliable predictor of company's future abnormal returns. At market level, we show that a long-short valuation (LSV) factor, built using buy-sell recommendations[1] based on the fair value distribution, contains information not accessible through traditional market factors. The LSV factor increases significantly the explanatory and predictive power of factor models estimated on portfolios and individual stocks returns.

4.1 Firm misvaluation z-score

Let μ_t^i and σ_t^i be the empirical mean and standard deviation of the log-fair value distribution of stock i, obtained from the bootstrapping procedure based on our SDCF method, at some time t, and let p_t^i be the closing log-price at day t of the same company. As mispricing indicator of company i at time t we take

$$z_t^i := \frac{p_t^i - \mu_t^i}{\sigma_t^i} \quad i \in \{1, \dots, N\} \ , \tag{4.1}$$

that is the log-difference of the company's expected fair value and its price, divided by the standard deviation of the log-fair value distribution. In our indicator, the absolute level of of misvaluation, $|p_t^i - \mu_t^i| = |p_t^j - \mu_t^j|$, is amplified when the uncertainty associated to the valuation procedure is less uncertain. Indeed, the uncertainty of the evaluation procedure, captured by the standard deviation σ_t^i, is used to modify the observed difference (in log) between the market price and the expected present value.

We expect an appropriate mispricing indicator to be related with future market adjustments, as prices of undervalued companies growth more than those of overvalued ones. With this hypothesis, we test the predictability power of our indicator with respect to observed future price returns.

[1] In Chapter 5 we will analyze how to build stock recommendation systems more specifically.

4.1.1 Analysis on the cross-section of individual stock returns

First, we assess whether the individual mispricing indicator z_t^i possesses significant predictability power for the one-quarter ahead excess return when used to augment factor models.

To this end, we perform panel fixed-effects regressions where all the cross-section of stocks excess returns are regressed on the z-scores and a set of control variables. For each month t, let $R_{i,t}^{\mathrm{EX}}$ be the monthly excess return of the firm i over the risk-free rate $R_{F,t}$, we consider the following model:

$$\begin{aligned} R_{i,t}^{\mathrm{EX}} &= \alpha_i + \beta_i^M \left(R_{M,t} - R_{F,t}\right) + \beta_i^{\mathrm{SMB}}\, SMB_t + \beta_i^{\mathrm{HML}}\, HML_t \\ &\quad + \gamma_1\, z_{t-3}^i + \gamma_2\, R_{t-1}^i + \gamma_3\, R_{t-12,t-2}^i + \gamma_4\, \log(ME_t^i) \\ &\quad + \gamma_5\, \log(BM_t^i) + \gamma_6\, ACC_t^i + \gamma_7\, AG_t^i + \gamma_8\, DE_t^i + e_{i,t}\,, \quad i \in \{1, \ldots, N\}, \end{aligned} \tag{4.2}$$

where $R_{M,t} - R_{F,t}$, SMB_t and HML_t are respectively the market factor, the size factor and the book-to-market factor of Fama-French three factor model; z_{t-3}^i is the z-score of the firm i computed averaging the daily z-scores in the previous quarter; R_{t-1}, $R_{t-12,t-2}$ are the past one-month return and the return from month $t-12$ to $t-2$; ME is the market equity; BM is the book-to-market ratio; ACC are the operating accruals; AG is the asset growth; DE is the leverage ratio and $e_{i,t}$ an idiosyncratic error term. Results are reported in Table 4.1 for seven different models with an increasing number of controls. The estimated γ's, i.e., the common effect of the mispricing score, are statistically significant and with a negative sign, irrespective of the number and type of control variables considered. In other terms, undervalued (resp. overvalued) stocks are, on average, consistently characterized by higher (resp. lower) future excess returns. This observation confirms the idea that our z-score represents a measure of misvaluation which is reabsorbed by the market over time, while price gradually converges to the company's fundamental value. Notice that each explanatory variable in (4.2) is cross-sectional normalized to have mean 0 and standard deviation 1. With some precautions due to to possible cross-correlation effects which might be neglected, this allows for a direct comparison of the regression coefficients. The picture that emerges from Table 4.1 is that, among all the considered regressors, the three main effects which seem to be more persistent are those of the z-score, the past one-month return and the book-to-market ratio.

4.1.2 Analysis at portfolio level

To further validate the capability of the z-score indicator to anticipate future market performances, we sort stocks into quantiles based on the z-score empirical distribution function at the beginning of each semester and then construct "Buy", "Hold" and "Sell" portfolios according to this quantile-based splitting. Specifically, let $\rho(\alpha)$ be the quantile function at level α of the empirical distribution of the z-scores z^i, $i \in \{1, \ldots, N\}$. If

Table 4.1 Results for the monthly fixed-effect time series regressions in (4.2). Coefficients resulting to be significant at 10%, 5%, 1% and 0.1% level are marked with '$\circ$', '$*$', '$**$' and '$***$' respectively. T-ratio based on HAC robust standard errors are reported in parentheses. All the control variables are provided by the Eikon database The sample period is April, 2009 to June, 2018. The regressors are (cross-sectional) centered to have mean 0 and (cross-sectional) scaled to have standard deviation 1.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
z-score	-0.280 ***	-0.331 ***	-0.500 ***	-0.485 ***	-0.503 ***	-0.475 ***	-0.473 ***
	[-3.318]	[-3.892]	[-6.134]	[-5.979]	[-6.159]	[-5.836]	[-5.797]
R_{t-1}		-0.460 ***	-0.500 ***	-0.524 ***	-0.500 ***	-0.519 ***	-0.518 ***
		[-3.539]	[-3.849]	[-4.006]	[-3.849]	[-3.915]	[-3.910]
$R_{t-12,t-2}$		-0.184	-0.200	-0.199	-0.200	-0.212	-0.207
		[-1.368]	[-1.486]	[-1.470]	[-1.488]	[-1.554]	[-1.518]
$\log(ME)$			-0.693 *	-0.730 *	-0.670 *	-0.764 **	-0.784 **
			[-2.523]	[-2.544]	[-2.438]	[-2.854]	[-2.777]
$\log(BM)$			1.174 ***	1.157 ***	1.163 ***	1.198 ***	1.207 ***
			[8.853]	[8.802]	[8.743]	[8.232]	[8.269]
ACC				0.176			0.203
				[1.248]			[1.444]
AG					-0.032		-0.027
					[-0.592]		[-0.506]
DE						-0.126 **	-0.129 **
						[-3.028]	[-3.112]
Adj. $R^2(\%)$	34.157	34.586	34.897	35.177	34.894	35.464	35.350
No. obs.	15540	15540	15429	15207	15429	14652	14541

$z^i < \rho(0.4)$ firm i is assigned to the Buy portfolio, if $z^i \geq \rho(0.6)$ firm i is assigned to the Sell portfolio and if $\rho(0.4) < z^i < \rho(0.6)$ it is assigned to the Hold portfolio. Notice that the Buy and Sell portfolios contain the same number of firms, while the Hold portfolios contains half that number of firms. For each portfolio we compute the equally weighted daily return and we compare its performance with the Our universe portfolio, defined as the equally weighted portfolio of all the stocks in our universe. Results are reported in Table 4.2. The Sharpe (Sharpe, 1994) and the Sortino (Sortino and Price, 1994) ratios associated with the Buy portfolio are 1.43 and 1.87 respectively, which are higher than those of the Sell, Hold and Our universe portfolios. The same conclusions hold for the average annual return. Using the test discussed in Ledoit and Wolf (2008) and Ardia and Boudt (2018), we found a significant difference between the Sharpe ratios of the Buy and the Our universe portfolios, with a *t-Statistic* of 2.98 and a *p-value* of $3e-03$. This cross-sectional investigation confirms the explanatory power of our misvaluation measure, as portfolios built using undervalued firms perform better than portfolios made of overvalued firms or made of all the firms in our reference universe.

In summary, the statistical analysis performed in this section reveal the capability

Table 4.2 Descriptive Statistics of Portfolios based on the z-scores measure. For each columns is reported the average number of firms, the average percentage of market capitalization respect to the our universe, the average annual log return, the annualized Sharpe and Sortino ratio, respectively.

	Sell	Hold	Buy	Our Universe
Avg. Number of Firms	56	28	56	140
Avg. Market Cap. (%)	34	16	50	100
Avg. Annual Return(%)	17.63	18.07	20.83	19.04
Sharpe Ratio	1.05	1.15	1.43	1.24
Sortino Ratio	1.36	1.48	1.87	1.59

of our misvaluation indicator of explaining a significant portion of the company future excess returns. A relevant question as this point in to understand how much of its predictive power our indicator retains when confronted with other possible sources of excess returns, as identified by the literature. We need to investigate if, and to what extent, the information revealed by our indicator represents a genuinely new contribution to the analysis of market dynamics which is not already contained in other variables the literature advanced as possible explanatory factors of stocks performances. This investigation is the focus of the next Section.

4.2 The valuation factor

To asses the predictive power of our misvaluation indicator with respect to future stocks performances we revert to the most traditional factor model analysis. We define a misvaluation factor LSV (Long Short Valuation) whose value at each day t is given by the difference between the equal-weighted return of a portfolio that goes long on the undervalued stocks and short on the overvalued ones. Therefore, the LSV factor is computed as the difference between the Buy and Sell portfolios discussed in Section 4.1.2. Over the sample period April 1, 2009 to September 28, 2018, this factor earns a mildly significantly positive average return of 2.7% ($t\text{-}Statistic = 1.49$ and $p\text{-}value = 0.14$) and has an annual Sharpe ratio of 0.48.

In the next section we compare LSV with other commonly considered factors affecting stocks return, namely the market factor, defined as the difference between the market return R_M and the risk-free interest rate R_F, the size factor (SMB), the book-to-market factor (HML), the momentum factor (MOM), the profitability (robust minus weak) factor (RMW) and the investment factor (CMA). See Fama and French (2015) and Carhart (1997) for a discussion on how these factor are built. In addition, we will investigate the relation of the LSV with the UMO factor, recently proposed in Hirshleifer and Jiang (2010) as a possible way of capturing the presence of a persistent, long-term, company misvaluation.

Table 4.3 Pearson correlation between the misvaluation factor *LSV* and other market factors. $R_M - R_F$ is the market factor, *SMB* the size factor , *HML* the book-to-market factor , *MOM* the momentum factor, *RMW* the profitability factor and *CMA* the investment factor and *UMO* is the financing-based misvaluation factor of Hirshleifer and Jiang (2010). The sample size is 2392 (reduced to 1953 for the *UMO* factro).

	LSV	$R_M - R_F$	SMB	HML	MOM	RMW	CMA	UMO
LSV	-	-0.405	-0.338	-0.035	-0.226	0.229	0.042	0.094
$R_M - R_F$		-	0.386	0.230	-0.120	-0.461	-0.038	-0.263
SMB			-	0.103	-0.099	-0.358	0.025	-0.320
HML				-	-0.421	-0.298	0.477	0.149
MOM					-	0.092	-0.069	0.039
RMW						-	-0.050	0.461
CMA							-	0.288
UMO								-

In Section 4.2.2 we will use the *LSV* to augment standard factor models and explore its relative merits using the Fama-MacBeth regression framework (Fama and French, 1992, Fama and MacBeth, 1973).

4.2.1 Comparing LSV with other market factors

We start with a simple bivariate analysis. Table 4.3 reports the Pearson correlation coefficient between *LSV* and the other considered factors over the sample period computed using daily returns. Our factor clearly shares some information content with the $R_M - R_F$, *SMB*, *MOM* and *RMW* factors. Conversely, its correlation with the *UMO* factor, over the entire period of analysis, does not seem significant. A visual inspection of the scatter plots, reported in Figure 4.1, confirms the apparent lack of correlation between *LSV* and *UMO* and a strong correlation of *LSV* with both $R_M - R_F$ and *SMB*. At this stage, the correlation with *MOM* seems instead due to few extreme observations (like the one with *RMW*, which is not reported for brevity).

The orthogonality of the *UMO* and *LSV* factors, emerging from Table 4.3 and confirmed in Figure 4.1, seems peculiar due to their shared claim of capturing the presence of market misvaluation. To qualify this finding, it is useful to compare the time profile of the two factors. In Figure 4.2 we plot the daily absolute value of the logarithmic return of the *UMO* and *LSV* factor rescaled by their mean and standard deviation. The absolute variation of the factor can be interpreted as a measure of its contribution in explaining market dynamics (Chang et al., 2013). As can be seen, the *UMO* factor identifies a high value of market misvaluation in the period between 2015 and 2016 while, according to the *LSV* factor, it is the window between 2010 and 2011 that display the most misvalued market prices. The two misvaluation factors are in some sense complementary and they seem to capture different phenomena. In fact, the presence of this difference in the time

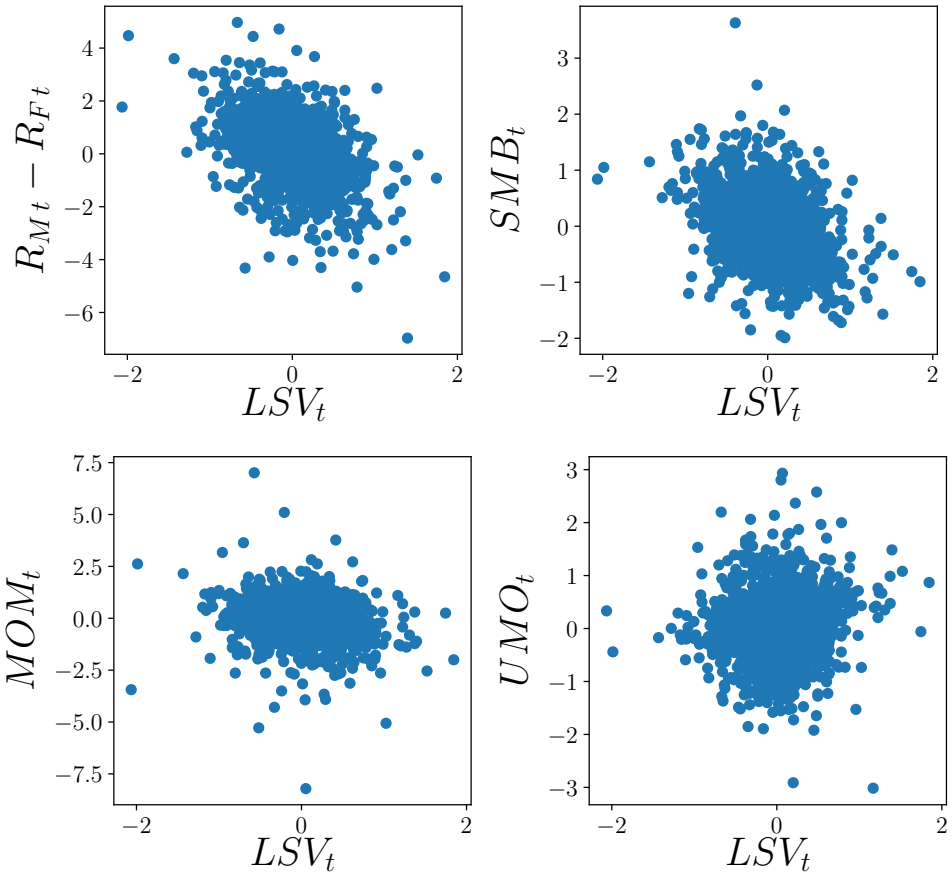


Fig. 4.1 Scatter plot between LSV_t and a selection of other factors.

structure of the two factors is not surprising if one looks at their definitions. The *UMO* factor is market-oriented and is based on stocks classification that looks to market operations (equity and debt offerings and buy-backs) in the previous two years, see Hirshleifer and Jiang (2010) for further details. Conversely, the *LSV* factor, based on the cross-section misvaluation indicators derived by the SDCF method, is more oriented toward company's operating performances and it is built starting from individual balance sheet data and revenues forecast. In the eve of the 2008 financial crisis, the scope of market operations was dramatically reduced, and the variability of the *UMO* factor consequently reduced. At the same time, the liquidity crises, especially in some sectors, induced significant misvaluation, whence the higher observed turbulence of the *LSV* factor in the years 2009-2010.

Next we move to a multivariate analysis regressing the *LSV* over the other considered factors. Consider the general model

$$\begin{aligned} LSV_t =& \beta_0 + \beta^{\mathrm{M}}(R_{M,t} - R_{F,t}) + \beta^{\mathrm{SMB}} SMB_t + \beta^{\mathrm{HML}} HML_t + \beta^{\mathrm{MOM}} MOM_t \\ &+ \beta^{\mathrm{RMW}} RMW_t + \beta^{\mathrm{CMA}} CMA_t + \beta^{UMO} UMO_t + e_t, \end{aligned} \tag{4.3}$$

where e_t is a zero-mean residual. The regressions results, in various model configurations,

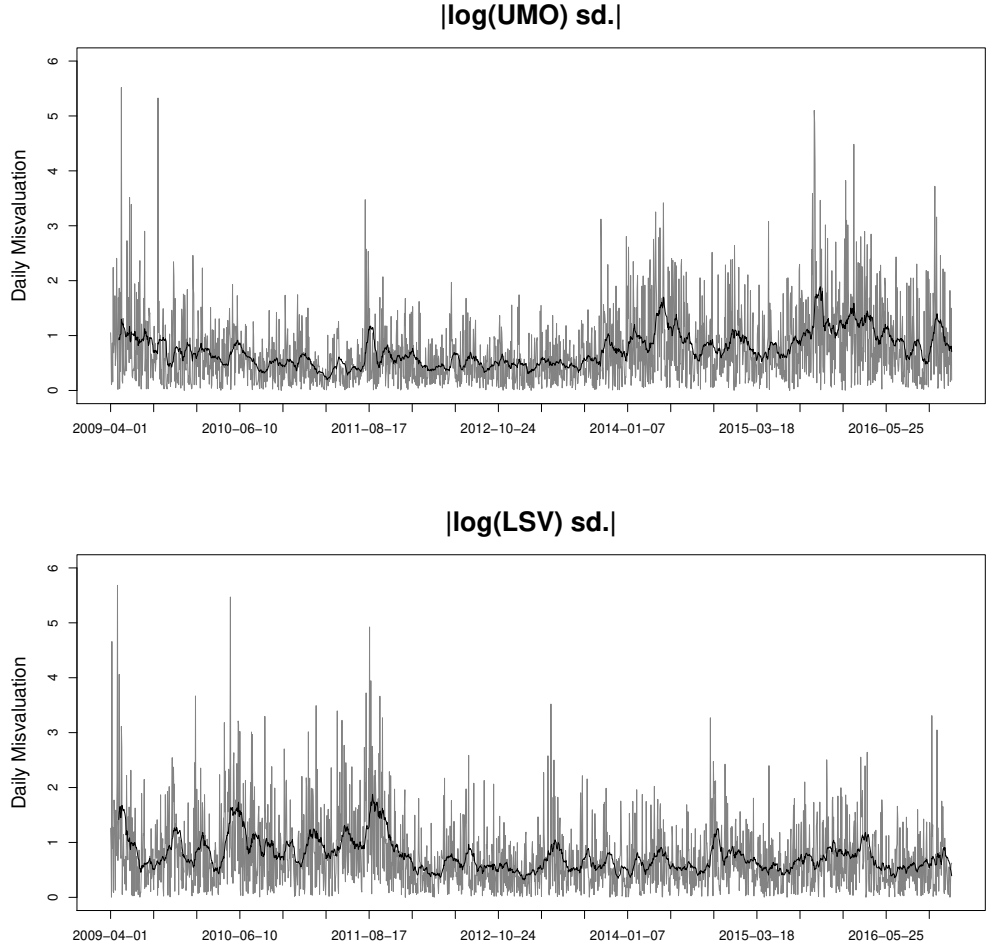


Fig. 4.2 Daily absolute value of the logarithmic return of the *UMO* and *LSV* factor rescaled by their mean and standard deviation. The dark solid lines are the moving averages over the past 20 days.

are displayed in Table 4.4. Both the value of the intercept, significantly different from zero at any conventional level in any setting, and the relatively small adjusted R^2, which is merely between 0.20 and 0.30 in all models, suggest that the *LSV* factor captures yet another anomaly which is hardly explained by well-accepted risk factors. The multivariate analysis confirms the correlation of *LSV* with the market, *SMB* and *MOM* factors. Conversely, *RMW*, *HML* and *CMA* results orthogonal to *LSV*. Notice that, when the interaction of different factors is taken into account, the correlation between *LSV* and *UMO* results negative and statistically significant at any conventional level. This is a further evidence of the different nature of the two factors.

4.2.2 The LSV beta and the Cross-Section of Portfolio (Abnormal) Returns

We now turn to our primary task in this section, testing, through factor model analysis, how well *LSV* explains average abnormal returns on the cross-section of portfolios. As

Table 4.4 Summary of the daily regressions of LSV on the factors $R_M - R_F$, SMB, HML, MOM, RMW, CMA and UMO. T-ratio based on HAC robust standard errors are reported in parentheses. Coefficients resulting to be significant at 10%, 5%, 1% and 0.1% level are marked with '$\circ$', '$*$', '$**$' and '$***$' respectively. The sample is from April 1, 2009 to September 28, 2018. When UMO factor is considered the data sample is from April 1, 2009 to December 30, 2016.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Intercept	0.021 **	0.020 **	0.021 **	0.021 **	0.020 **	0.025 ***	0.022 **
	[3.160]	[3.269]	[3.115]	[3.168]	[3.242]	[3.303]	[3.240]
$R_M - R_F$	-0.125 ***	-0.122 ***	-0.122 ***	-0.126 ***	-0.117 ***	-0.132 ***	-0.120 ***
	[-8.286]	[-10.043]	[-7.843]	[-8.087]	[-9.316]	[-9.040]	[-9.082]
SMB	-0.141 ***	-0.155 ***	-0.137 ***	-0.141 ***	-0.156 ***	-0.178 ***	-0.187 ***
	[-6.716]	[-8.927]	[-6.572]	[-6.795]	[-8.941]	[-7.912]	[-9.604]
HML	0.059 *	-0.025	0.062 *	0.063 $\circ$	-0.042	0.075 **	-0.020
	[2.376]	[-1.213]	[2.368]	[1.955]	[-1.636]	[2.704]	[-0.602]
MOM		-0.135 ***			-0.138 ***		-0.134 ***
		[-7.935]			[-8.045]		[-7.380]
RMW			0.026		0.005		0.070 $\circ$
			[0.933]		[0.214]		[1.934]
CMA				-0.015	0.056		0.058
				[-0.303]	[1.462]		[1.274]
UMO						-0.069 ***	-0.072 ***
						[-3.656]	[-3.679]
Adj. R^2(%)	20.912	28.561	20.928	20.890	28.655	23.661	31.192
No. obs.	2392	2392	2392	2392	2392	1953	1953

observed by Hirshleifer and Jiang (2010) for the UMO factors, we expect to obtain more stable loadings on portfolios that are formed based on possible measures of mispricing. Therefore, we select the Fama-French 25 portfolios formed on size and book-to-market and we examine the effect of LSV factor and other market factors by computing the average premium using Fama-MacBeth regression framework, see, e.g., Fama and MacBeth (1973), Fama and French (1992). A positive relation between abnormal returns and factor loadings suggests the existence of a systematic stock misvaluation positively captured by the LSV factor. In other terms, a positive (negative) loading on LSV indicates a systematic undervaluation (overvaluation) (Chang et al., 2013, Hirshleifer and Jiang, 2010).

We want to analyze both the explanatory and predictive power of LSV loadings on the cross-section of portfolios abnormal returns and therefore we investigate this relation in an in-sample and out-of-sample setting,[2]. Table 4.5 exhibits in-sample Fama-MacBeth

[2]We also perform the following analysis using directly the portfolios excess returns. We find similar results for the in-sample setting, which are reported in Appendix B, even if the statistics are slightly deteriorated, while for the out-of sample case we did not find a sufficient statistical evidence for all factors premia and they are available upon request.

Table 4.5 Fama-MacBeth monthly regressions on the cross-section of portfolio abnormal returns. The dependent variables are the monthly abnormal returns of 25 portfolios based on size and book-to-market, which are computed using Fama and French three-factor model on equally-weighted excess returns of portfolios. The independent variables are the loadings on the factors LSV, $LSV^{\perp}$ $R_M - R_F$, SMB, HML, MOM, RMW, CMA and UMO. T-ratio based on robust standard errors to HAC are reported in parentheses. Coefficients resulting to be significant at 10%, 5%, 1% and 0.1% level are marked with '$\circ$', '$*$', '$**$' and '$***$' respectively. The sample is from April 2009 to September 2018. When UMO factor is considered the data sample is from April 2009 to December 2016.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Intercept	0.330	1.132 *	1.765 ***	1.838 ***	1.863 ***	2.228 ***	2.241 ***
	[1.510]	[2.683]	[5.035]	[5.385]	[4.862]	[6.262]	[6.270]
LSV	0.313		0.492 *	0.747 **		0.762 **	
	[1.608]		[2.273]	[3.346]		[3.729]	
$LSV^{\perp}$							0.543 *
							[2.812]
$R_M - R_F$		-1.015 *	-1.812 ***	-1.750 ***	-1.897 ***	-2.180 ***	-2.193 ***
		[-2.618]	[-5.435]	[-5.362]	[-5.081]	[-6.466]	[-6.479]
SMB		-0.153	-0.062	-0.001	-0.033	-0.037	-0.036
		[-1.331]	[-0.704]	[-0.014]	[-0.333]	[-0.441]	[-0.424]
HML		0.136	-0.001	-0.026	0.056	-0.002	-0.001
		[1.221]	[-0.009]	[-0.274]	[0.543]	[-0.018]	[-0.006]
MOM			-0.536		-0.618	-0.449	-0.446
			[-1.135]		[-1.246]	[-1.046]	[-1.037]
RMW			0.808 ***		0.568 **	0.611 ***	0.606 ***
			[5.427]		[3.281]	[4.091]	[4.069]
CMA			-0.640 **		-0.345	-0.355 *	-0.353 $\circ$
			[-3.096]		[-1.593]	[-1.911]	[-1.904]
UMO				1.423 ***	1.104 ***	1.315 ***	1.319 ***
				[6.397]	[4.537]	[5.895]	[5.903]
Adj. R^2(%)	6.203	25.957	65.719	65.219	63.382	73.090	73.116

results based on monthly abnormal return, computed using the Fama and French three-factor model, of the 25 size-BM portfolios. In addition to the LSV, we consider the five traditional Fama and French, the momentum and the UMO factors as potential confounding explanatory variables. The results of the in-sample analysis are reported in Table 4.5. As expected, from the nature of the considered portfolios, the SMB and HML factors are constantly not-significant. The monthly average premium of the LSV factor is always positive and significantly different from zero when all factors are considered, see Columns (3) - (6). Remarkably, this remains true also for the model in Column (7), where we consider an orthogonalized misvaluation factor, $LSV^{\perp}$, defined as the sum of the intercept and residuals extracted from the regression in (4.3), i.e., $LSV^{\perp} := \beta_0 + e_t$. By construction, the orthogonalized misvaluation factor has zero correlation with the Fama-French, MOM and UMO factors. Notice that the loadings of the UMO are concordant

Table 4.6 Out of sample Fama-MacBeth daily regressions on the cross-section of portfolio abnormal returns. The dependent variables are the abnormal returns of 25 portfolios based on size and book-to-market, which are computed using Fama and French three-factor model on equally-weighted excess returns of portfolios. The independent variables are the loadings on the factors LSV, $LSV^{\perp}$ $R_M - R_F$, SMB, HML, MOM, RMW, CMA and UMO. The time-series averages of the cross-sectional regression coefficients are reported. The Avg. R^2's are the time series averages of the adjusted R^2 across the full sample period. T-ratio based on robust standard errors to HAC are reported in parentheses. Coefficients resulting to be significant at 10%, 5%, 1% and 0.1% level are marked with '$\circ$', '$*$', '$**$' and '$***$' respectively. The sample is from April 1, 2009 to September 28, 2018. When UMO factor is considered the data sample is from April 1, 2009 to December 30, 2016.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Intercept	0.027 **	0.083 ***	0.077 ***	0.094 ***	0.084 ***	0.079 ***	0.079 ***
	[2.968]	[4.446]	[4.440]	[4.603]	[4.217]	[3.916]	[3.916]
LSV	0.011 $\circ$		0.017 $\circ$	0.014		0.028 *	
	[1.857]		[1.711]	[1.481]		[2.537]	
$LSV^{\perp}$							0.016
							[1.394]
$R_M - R_F$		-0.080 ***	-0.078 ***	-0.091 ***	-0.085 ***	-0.080 ***	-0.080 ***
		[-4.350]	[-4.487]	[-4.436]	[-4.159]	[-3.903]	[-3.903]
SMB		-0.001	0.004	-0.001	0.004	0.004	0.004
		[-0.317]	[1.317]	[-0.204]	[0.832]	[0.982]	[0.982]
HML		0.005	0.001	0.003	0.006	0.003	0.003
		[1.256]	[0.364]	[0.581]	[1.262]	[0.744]	[0.744]
MOM			-0.014		-0.020	-0.019	-0.019
			[-0.779]		[-1.059]	[-0.920]	[-0.920]
RMW			0.014 $\circ$		0.010	0.012	0.012
			[1.938]		[1.108]	[1.190]	[1.190]
CMA			0.011		0.021 *	0.024 **	0.024 **
			[1.437]		[2.279]	[2.667]	[2.667]
UMO				0.023 *	0.024 $\circ$	0.026 $\circ$	0.026 $\circ$
				[2.118]	[1.820]	[1.885]	[1.885]
Avg. $R^2(\%)$	4.807	16.568	29.480	28.953	32.993	34.273	34.273

with the loadings of the LSV and $LSV^{\perp}$ factors. The asset "misvaluation" the two factors are built to capture, while different in nature, is still consistent in predicting higher (lower) returns for undervalued (overvalues) stocks.

Then we move to an out-of-sample analysis using a 60 days rolling window updated every 30 days: for each portfolio at each date, the loadings on the considered factors are estimated from a time-series regression using daily excess returns over the previous 60 days. Then, the future abnormal returns of each portfolio are computed by regressing the equally-weighted excess returns on the Fama and French three-factor model over the following 30 days. The estimated abnormal returns and factor loadings are then used

as dependent and independent variables, respectively, in the cross-sectional regressions. Table 4.6 reports the average premia of the out-of-sample analysis and the related statistics. The market factor of Fama and French is the only factor which possess a strongly significant premium for all model specifications. However, we find that the LSV premium is always positive and it results significant when all other factor are considered, column (6).

4.2.3 The LSV beta and the Cross-Section of Individual Stock Returns

As highlighted by Hirshleifer and Jiang (2010) and Chang et al. (2013) factor loadings on individual stocks tends to be unstable. Therefore, the analysis of a possible relation between LSV loadings on individual stocks can be "challenging", Hirshleifer and Jiang (2010). We follow their approach to test for stability, but in contrast to them, we examine the $LSV^{\perp}$ loadings (instead of those of LSV) in order to study the unique information contained in the LSV factor, since it is obtained by removing all possible collinear information of all other market factors. We estimate $LSV^{\perp}$ betas from daily excess returns using the following model over 100 days for each firm i of the S&P500:

$$\begin{aligned} R_{i,t} - R_{F,t} = \alpha_i + \beta_i^M \left(R_{M,t} - R_{F,t}\right) + \beta_i^{SMB} SMB_t + \beta_i^{HML} HML_t + \\ + \beta_i^{UMO} UMO_t + \beta_i^{LSV} LSV_t^{\perp}. \end{aligned} \tag{4.4}$$

We then use deciles based on β^{LSV} to sort firms and for each decile we evaluate the performance on the succeeding 30 days. Table 4.7 reports the annual average equally-weighted decile returns together with the related abnormal returns computed using $CAPM$, Fama and French three-factor model and Carhart four-factor model where each model is augmented with UMO factor, respectively.

We also report the corresponding performance of the returns of all equally-weighted stocks, i.e., when there is no sorting, which is used as benchmark. We denote this class as *All*. We observe that decile returns tend to increase with the ranking based on β^{LSV}. In particular, the statistical significance seems to be an increasing function of decile ranking class, which suggests that the high ranking classes are less volatile and then with a more robust performance, e.g., the Higher (H) and Lower (L) classes have a annual returns equal to 15.796% (t-stat =2.583) and 14.615% (t-stat =1.798), respectively[3]. This is more evident when we consider the corresponding abnormal returns. Indeed, only the high three classes earn a statistically positive abnormal return, which are also greater than the benchmark abnormal return of all equally-weighted stocks, e.g., the annual percentage abnormal return computed using Carhart four-factor model augmented with

[3]Even if, some lower classes earn average returns which seem to be greater than that of the corresponding higher classes, by performing a one-way ANOVA we find no statistical evidence (F-stat=0.02) among average returns, which points out how the difference between the deciles is in terms of more robust performance, i.e., more robust t-statistics.

Table 4.7 Performance of deciles based on $LSV^{\perp}$ loadings. This table reports the average percentage annual equally-weighted returns and the corresponding average percentage annual abnormal returns for each decile based on the $LSV^{\perp}$. The loadings are computed over 100 days and firms are sorted in deciles for the next 30 days. The post-ranking β^{LSV}_{post} loadings are estimated using all the full sample decile returns. Abnormal returns are computed using $CAPM$ model, Fama and French three-factor model and Carhart four-factor model augmented with UMO factor, respectively. The H-L row corresponds to a portfolio which is long on the higher (H) ranking and short to the lower (L) one based on β^{LSV}_{pre} loadings. The last row (All) is the performance when all deciles are merged. T-ratio based on robust standard errors to HAC are reported in parentheses. Coefficients resulting to be significant at 10%, 5%, 1% and 0.1% level are marked with '$\circ$', '$*$', '$**$' and '$***$' respectively. The data sample is constituted by the S&P500 firms (at April 2009) and start from April 1, 2009 to 30 December 2016.

	β^{LSV}_{pre}	Ret	β^{LSV}_{post}	$\alpha_{CAPM+UMO}$	$\alpha_{FF3+UMO}$	$\alpha_{FF3+MOM+UMO}$
L	-1.436	14.615 $\circ$	-0.666	-0.698	-0.643	-0.562
		[1.798]		[-0.253]	[-0.247]	[-0.216]
2	-0.759	17.296 *	-0.406	2.166	2.284	2.310
		[2.377]		[1.158]	[1.290]	[1.299]
3	-0.475	15.578 *	-0.252	0.743	0.845	0.869
		[2.289]		[0.507]	[0.570]	[0.585]
4	-0.268	14.332 *	-0.086	0.296	0.223	0.211
		[2.251]		[0.207]	[0.157]	[0.148]
5	-0.100	14.769 *	-0.030	0.938	1.022	1.027
		[2.425]		[0.692]	[0.763]	[0.769]
6	0.053	15.443 **	0.021	2.316	2.331	2.315
		[2.643]		[1.625]	[1.624]	[1.618]
7	0.209	14.352 *	0.066	1.456	1.436	1.401
		[2.567]		[1.078]	[1.059]	[1.039]
8	0.380	15.378 **	0.122	2.876 $\circ$	2.884 $\circ$	2.840 $\circ$
		[2.852]		[1.862]	[1.857]	[1.848]
9	0.601	15.679 **	0.206	3.850 *	3.861 *	3.791 *
		[2.940]		[2.044]	[2.043]	[2.017]
H	1.155	15.796 **	0.412	4.246 $\circ$	4.436 $\circ$	4.493 $\circ$
		[2.583]		[1.656]	[1.682]	[1.693]
H-L	2.591	1.181	1.082	6.605 $\circ$	6.738 $\circ$	6.715 $\circ$
		[0.256]		[1.653]	[1.648]	[1.697]
All	-0.062	15.327 *	-0.061	1.708 *	1.757 *	1.758 *
		[2.511]		[2.104]	[2.242]	[2.215]

UMO is 4.493% and 1.758% for H and All classes, respectively. Moreover, we observe that abnormal returns remain stable and statistically significant among the $CAPM$, Fama and French three-factor and Carhart four-factor model (augmented with UMO). In Table 4.7 it is also reported, for each decile, the post ranking $LSV^{\perp}$ loadings, which are computed using the same model (4.4) over the full sample decile returns. We denote

as β_{pre}^{LSV} and β_{post}^{LSV} the initial and post ranking $LSV^{\perp}$ loadings, respectively. The β_{post}^{LSV} loadings increase with the corresponding β_{pre}^{LSV}, which suggests that there is a persistence among the $LSV^{\perp}$ loadings over a 30 days window. Furthermore, in Table 4.8 we also replicate the previous analysis excluding the firms of our Universe, which are employed for the construction of the LSV factor, and we observe similar results which confirms the significant role of LSV loadings on individual stocks returns.

In conclusion, the analyses of this Chapter reveal the presence of relevant information captured by the LSV factor which is complementary with respect to the information made available by other market factors. This is more evident at firm level than in the portfolio aggregate, even if we find a significant positive explanatory relation between LSV loadings and portfolio abnormal returns. These results confirm the relevance of the misvaluation indicator derived via the SDCF method starting from individual firm balance sheet data, both as a genuinely new explanatory variable of contemporaneous price movements, and as a potential predictor of future market performances.

Table 4.8 Performance of deciles based on $LSV^{\perp}$ loadings when "Our universe" of stocks is excluded. This table reports the average percentage annual equally-weighted returns and the corresponding average percentage annual abnormal returns for each decile based on the $LSV^{\perp}$. The loadings are computed over 100 days and firms are sorted in deciles for the next 30 days. The post-ranking β^{LSV}_{post} loadings are estimated using all the full sample decile returns. Abnormal returns are computed using $CAPM$ model, Fama and French three-factor model and Carhart four-factor model augmented with UMO factor, respectively. The H-L row corresponds to a portfolio which is long on the higher (H) ranking and short to the lower (L) one based on β^{LSV}_{pre} loadings. The last row (All) is the performance when all deciles are merged. T-ratio based on robust standard errors to HAC are reported in parentheses. Coefficients resulting to be significant at 10%, 5%, 1% and 0.1% level are marked with '$\circ$', '$*$', '$**$' and '$***$' respectively. The data sample is constituted by the S&P500 firms (at April 2009), when all firms employed for the LSV construction are removed, and start from April 1, 2009 to 30 December 2016.

	β^{LSV}_{pre}	Ret	β^{LSV}_{post}	$\alpha_{CAPM+UMO}$	$\alpha_{FF3+UMO}$	$\alpha_{FF3+MOM+UMO}$
L	-1.388	16.204 $\circ$	-0.564	1.161	1.380	1.480
		[1.948]		[0.373]	[0.467]	[0.502]
2	-0.721	17.770 *	-0.310	2.563	2.896	2.935
		[2.403]		[1.201]	[1.449]	[1.461]
3	-0.448	15.168 *	-0.164	0.362	0.663	0.678
		[2.186]		[0.200]	[0.373]	[0.378]
4	-0.252	14.270 *	-0.059	0.186	0.271	0.254
		[2.195]		[0.109]	[0.159]	[0.149]
5	-0.089	15.724 *	-0.019	1.871	2.043	2.035
		[2.498]		[1.219]	[1.344]	[1.339]
6	0.055	14.026 *	0.014	0.760	1.024	0.996
		[2.313]		[0.442]	[0.598]	[0.583]
7	0.204	14.267 *	0.044	1.117	1.212	1.166
		[2.499]		[0.697]	[0.754]	[0.732]
8	0.373	14.061 *	0.071	1.417	1.567	1.506
		[2.484]		[0.756]	[0.827]	[0.804]
9	0.595	16.743 **	0.150	4.705 *	4.904 *	4.817 *
		[3.002]		[2.170]	[2.256]	[2.236]
H	1.112	17.703 **	0.277	5.985 *	6.105 *	6.116 *
		[2.727]		[2.073]	[2.075]	[2.072]
H-L	2.499	1.499	0.845	6.463	6.363	6.274
		[0.312]		[1.520]	[1.506]	[1.491]
All	-0.056	15.593 *	-0.056	1.870 $\circ$	2.064 *	2.056 *
		[2.488]		[1.913]	[2.207]	[2.184]

Appendix B

In-sample analysis of LSV beta and excess returns

In Table B.1 we report the in-sample results on the cros-section of portfolio returns, discussed in Section 4.2.2, to test the explanatory power on the excess returns. The analogues columns of (1) and (2) of Table 4.5 are omitted since all premia are not significant for all factors.

Table B.1 Fama-MacBeth monthly regressions on the cross-section of portfolio returns. The dependent variables are the monthly excess returns of 25 portfolios based on size and book-to-market. The independent variables are the loadings on the factors LSV, $LSV^{\perp}$ $R_M - R_F$, SMB, HML, MOM, RMW, CMA and UMO. T-ratio based on robust standard errors to HAC are reported in parentheses. Coefficients resulting to be significant at 10%, 5%, 1% and 0.1% level are marked with '$\circ$', '$*$', '$**$' and '$***$' respectively. The sample is from April 2009 to September 2018. When UMO factor is considered the data sample is from April 2009 to December 2016.

	(1)	(2)	(3)	(4)	(5)
Intercept	1.765 ***	1.700 ***	1.806 ***	2.136 ***	2.150 ***
	[5.035]	[4.916]	[4.882]	[6.055]	[6.076]
LSV	0.280	0.491 *		0.506 *	
	[1.295]	[2.168]		[2.499]	
$LSV^{\perp}$					0.417 *
					[2.180]
MKT	-0.430	-0.280	-0.510	-0.765 *	-0.779 *
	[-1.289]	[-0.847]	[-1.414]	[-2.289]	[-2.325]
SMB	0.106	0.190 $\circ$	0.155	0.151 $\circ$	0.152 $\circ$
	[1.202]	[1.968]	[1.631]	[1.804]	[1.824]
HML	-0.060	-0.082	0.003	-0.049	-0.048
	[-0.642]	[-0.855]	[0.034]	[-0.540]	[-0.534]
MOM	-0.834 $\circ$		-0.954	-0.802 $\circ$	-0.798 $\circ$
	[-1.768]		[-1.993]	[-1.883]	[-1.877]
RMW	0.665 ***		0.341	0.380 *	0.376 *
	[4.466]		[2.040]	[2.565]	[2.548]
CMA	-0.701 **		-0.419	-0.427 *	-0.426 *
	[-3.389]		[-2.003]	[-2.322]	[-2.320]
UMO		0.918 ***	0.654 *	0.844 **	0.849 **
		[4.071]	[2.783]	[3.819]	[3.840]
Adj. $R^2(\%)$	50.216	41.510	44.131	56.689	56.894

CHAPTER 5

Stock Recommendations from Stochastic Discounted Cash Flows

We present two stock recommendation systems based on stochastic characterization of the firm present value. The *single-stock quantile* recommendation system compares the market price of a company's stocks with the estimated distribution of the company's fair value to obtain an individual measure of mispricing, while the *cross-sectional quantile* system builds a relative measure of mispricing using the fair-value distribution of all firms. Both systems use mispricing information to build sell side and buy side portfolios. Statistical tests show that these portfolios consistently deliver significant excess returns even when rebalancing costs are accounted for. Moreover, we show how analysts' indications can be improved by using the first two moments of their recommendations following the same procedure of the cross-sectional quantile system.

5.1 Recommendations from fair-value distributions

The fair value distribution defined in the Chapter 3 can be straightforwardly used to obtain portfolio recommendations for company stocks. The basic idea is to use the valuation model to identify mispriced companies. Under the hypothesis that mispriced companies will revert to their correct price, undervalued firms represent prospective buys and overvalued firms represent prospective sells. Following standard practice, stocks are classified as strong buy (SB), buy (B), hold (H), sell (S) and strong sell (SS).

5.1.1 Single-Stock Quantile recommendations system

Let FV_t^i be the distribution function of the fair value of company i at time t and P_t^i be its market price. The quantity $q_t^i = FV_t^i(P_t^i)$ represents the probability that the company's fair value is less than or equal to the observed price. In general, if q_t^i is near 0.5, the market price is near the median of the fair value distribution and we can conclude that, the company is fairly priced. However, if the value assigned by our valuation model to the company is higher (lower) than the market price, then the company is undervalued (overvalued) and q_t^i is close to zero (one). Based on this consideration, the classification

of stocks is performed in the following way[1]:

- if $q_t^i < 0.125$, the company i is classified SB;
- if $0.125 \leq q_t^i < 0.25$ it is classified B;
- if $0.25 \leq q_t^i < 0.75$ it is classified H;
- if $0.75 \leq q_t^i < 1$ it is classified S;
- and SS if $1 \leq q_t^i$.

This classification system, denoted as *Single-Stock Quantile* (SSQ), has the advantage of using all the information provided by the distribution of the company's fair value. The recommendation for each firm is obtained using only its own fair value distribution, without referring to the valuation of other firms. Thus, it is possible that some of the recommendation buckets remain empty; for example, that no stock is labeled SB or SS.

5.1.2 Cross-Sectional Quantile recommendations system

A second approach is to use the fair value distribution of all firms at the same time. For this, we introduce a second recommendation system based on the definition of the company specific mispricing indicator z-score introduced in Section 4.1. Now, consider the empirical distribution function of all mispricing indicators z_t^i, $\forall i = 1, \ldots, N$ and let $\rho_t(\alpha)$ be its α-quantile:

- The sock of company i is classified SB if $z_t^i < \rho_t(0.1)$;
- B, if $\rho_t(0.1) \leq z_t^i < \rho_t(0.4)$;
- H, if $\rho_t(0.4) \leq z_t^i < \rho_t(0.6)$
- S, if $\rho_t(0.6) \leq z_t^i < \rho_t(0.9)$;
- and SS if $\rho_t(0.9) \leq z_t^i$.

Firms with a misvaluation indicator near the median of the empirical distribution of all indicators are assigned to the hold class. Firms with a high mispricing with respect to the median are assigned to the sell class and become a strong sell if they are in the top decile. Conversely, firms with a low misvaluation with respect to the median are a buy and a strong buy if they are in the bottom decile. We term this system *Cross-Sectional Quantile* (CSQ). The advantage of this system is that an overall shift in market prices that has no effect on the relative rankings of different companies has no effect on their

[1]The class assignment is broadly in line with the values adopted by the *Morningstar*® equity research methodology, see MorningstarEquityResearchMethodology.pdf for further details.

classification. This system is also insensitive to the presence of a common bias affecting the valuation procedure of different companies.

To test the performance of the SSQ and CSQ systems, we consider 19 non-overlapping periods of six months, from FQ1 2009 to FQ1 2018[2]. At the beginning of each period, we classify the firms using both systems. We use the first available closing price for computing the mispricing indicator. In the case of SSQ, both the number of firms in each class and the associated market capitalization, with respect to our universe of stocks, can vary from period to period, while for CSQ, the number of stocks in each class is constant in all periods. On average, in the SSQ system, the SB class has 31 stocks (28% market capitalization), B has 25 (18%), H has 62 (40%), S has 18 (11%) and SS has 3 (3%); while in the CSQ system, SB class has 14 stocks (13% market capitalization), B has 42 (37%), H has 28 (16%), S has 42 (25%) and SS has 14 (9%).

For each recommendation system, we build equally weighted portfolios with all companies in a given rating class at the beginning of each semester and compute the daily returns of these portfolios R_t^p, with p taking values SS, S, H, B, SB, on each day t of the semester[3].

5.2 Performance evaluation

We begin with a simple calculation, over the entire period considered, of the annualized log-returns (as percentages) for each of our constructed portfolios. In Figure 5.1 they are compared with the annualized log-returns of a benchmark equally weighted portfolio that goes long in all the stocks of our universe, labeled *Our universe*. As can be seen, undervalued assets tend to grow significantly faster than overvalued ones. For instance, the annualized log-return of the *SB* portfolio built following the *SSQ* system is 22.00% while that of the *SS* portfolio is 15.20%. This also holds true if we consider a more coarse grained classification, merging portfolios in the buy and sell sides. The use of just two broad classes seems to enhance the performance of the CSQ system. The enhanced performance of the portfolios obtained with the two recommendation systems is also confirmed when a measure of risk is included. The Sharpe Ratio (Sharpe, 1994) of the *Buy Side* portfolios is 1.43 for the CSQ system and 1.40 for the SSQ system. They are both significantly higher, according to the Ledoit and Wolf (2008) and Ardia and Boudt (2018) tests (p-value around 0.003), than the Sharpe Ratio of the *Our universe* portfolio, which is 1.24. In turn, the two sell side portfolios have values that are significantly lower

[2]The period of six months was chosen because it is long enough for the calibration of the cash flow model to be reliable but short enough to give us a sufficient number of data points to analyze. In any case, it is broadly consistent with several portfolio strategies discussed in (Li et al., 2019).

[3]In Barber et al. (2001), market-weighted rather than equally weighted portfolios are considered. Their choice is consistent with the use of daily rebalancing and the size of their sample. However, the authors warn about the possibility that using market-weighted returns could bias against finding evidence of abnormal returns, so we opt for a more conservative choice.

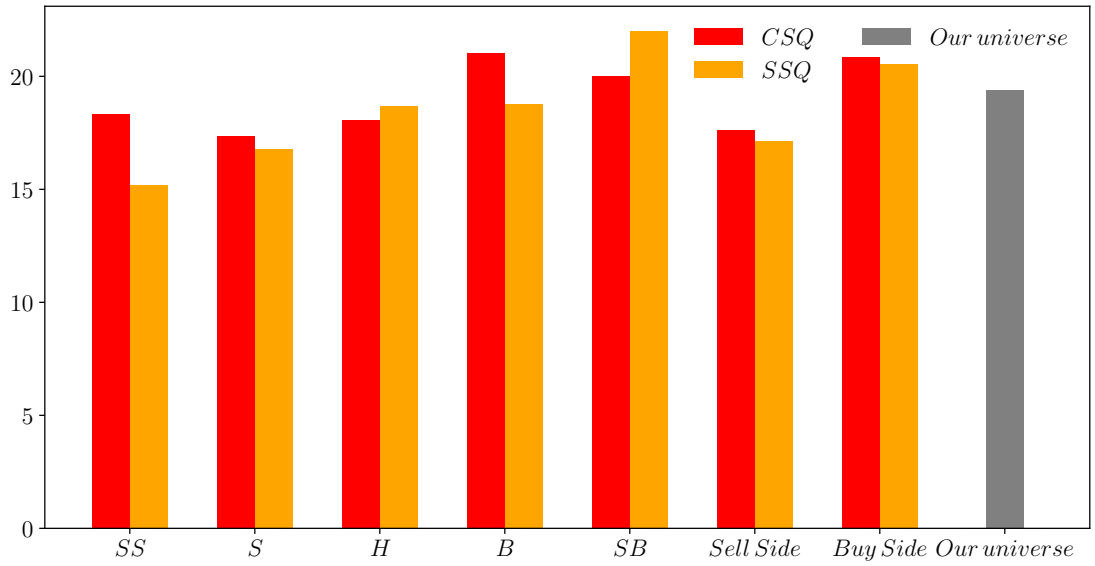


Fig. 5.1 Annualized log-returns in percentage for each of the constructed portfolio according to *CSQ* and to the *SSQ* recommendation. *Our universe* is the return of an equally-weighted portfolio that goes long in all the stocks of our universe. The sample period is April 1, 2009 to September 28, 2018.

than the Our Universe portfolio (1.05 for the *CSQ* system and 0.97 for the *SSQ* system)[4]. Figure 5.2 represents the cumulative sum of daily returns when investing \$100 in the *Buy Side*, *H* and *Sell Side* portfolios constructed with the *CSQ* (left panel) and *SSQ* (right panel) methodologies, compared with the *Our universe* portfolio.

To obtain a more precise estimate of portfolio performances, we employ an intercept test using the Fama-French three-factor model (Fama and French, 1993), augmented with the momentum factor (Carhart, 1997). We estimate the following daily time-series regression:

$$\begin{aligned} R_{p,t} - R_{F,t} = \alpha_p + \beta_p^M (R_{M,t} - R_{F,t}) \\ + \beta_p^{SMB} SMB_t + \beta_p^{HML} HML_t + \beta_p^{MOM} MOM_t + e_{p,t}\,, \end{aligned} \tag{5.1}$$

where $R_{p,t} - R_{F,t}$ denotes the excess return of the selected portfolio over the risk-free rate $R_{F,t}$ for period t, $R_{M,t}$ is the return of the value-weighted market portfolio, SMB_t is the difference between the daily returns of a value-weighted portfolio of small stocks and one of large stocks, HML_t is the difference between the daily returns of a value-weighted portfolio of high book-to-market stocks, MOM_t is the momentum factor and $e_{p,t}$ is the error term. The regression yields parameter estimates of α_p, β_p^M, β_p^{SMB}, β_p^{HML} and β_p^{MOM} but the relevant parameter here is intercept α_p, as it captures the presence of

[4] The Sharpe ratio is computed by setting the benchmark return to zero. We compared the portfolios performances using the Sortino ratio (Sortino and Price, 1994) also, obtaining identical results.

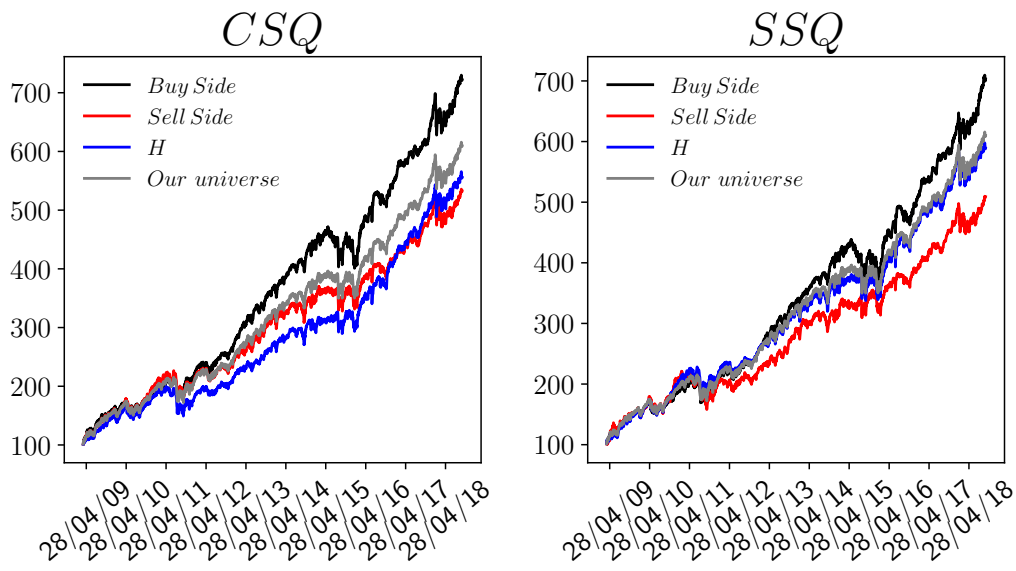


Fig. 5.2 Cumulative sum of daily returns when investing \$100 in the *Buy Side*, *H* and *Sell Side* portfolio constructed with the *CSQ* (left panel) and *SSQ* (right panel) methodologies. The grey line is the cumulative sum of daily returns of *Our universe*. The sample period is from April 1, 2009 to September 28, 2018.

Table 5.1 Estimated gross annual abnormal returns earned by portfolios constructed with our *CSQ* and *SSQ* systems, and using the analysts' recommendation from the I/B/E/S database. The coefficients significant at 10%, 5%, 1% and 0.1% level are marked with '$\circ$', '$*$', '$**$' and '$***$' respectively. The *SS* portfolios in the case of *SSQ* system and both the *SS* and *BB* portfolios in the case of analysts are empty for a few rebalancing dates. In such cases, the corresponding returns are set to zero.

α_p	*SS*	*S*	*H*	*B*	*SB*	*Sell Side*	*Buy Side*
CSQ	1.5733	0.8190	2.4977	5.7706 ***	7.0838 ***	1.0076	6.0989 ***
SSQ	1.4665	1.2052	2.0189	3.3869 *	7.9444 ***	1.4182	5.8442 ***
Analysts	11.0480 **	5.0222 **	2.4576 *	2.1187	1.9982	5.795 ***	2.0579

abnormal returns. The results are reported in Table 5.1. The most highly recommended stocks (*B*, *SB*, and *Buy Side*) earn positive abnormal gross returns, whereas the least recommended ones do not. In addition, the abnormal gross excess returns of these portfolios are greater than that of *Our Universe*, which has a gross annual excess return of 3.34% with a p-value of $3.16e-04$. These results suggest that investors following our SDCF based recommendations and building concentrated portfolios could obtain returns that beat the market.

As a further check, we repeat the same analysis using expert recommendations from the I/B/E/S database. Let $\overline{A}_t^i$ be the average analysts rating for firm i on date t. We follow Barber et al. (2001) and if $1 \leq \overline{A}_t^i \leq 1.5$ we classify company i as *SB*; if $1.5 < \overline{A}_t^i \leq 2$

Table 5.2 This table reports the Sharpe and the Sortino ratios, in parenthesis, for the *H*, *Sell Side* and *Buy Side* portfolios constructed according to *Analysts*, *Cross-Sectional Quantile* and *Single-Stock Quantile*. Sharpe and the Sortino ratios for *Our universe* is also reported.

Sharpe-ratio (Sortino-ratio)	*Analysts*	*CSQ*	*SSQ*	*Our universe*
Sell Side	1.31 (1.72)	1.05 (1.36)	0.97 (1.28)	
H	1.18 (1.52)	1.15 (1.48)	1.11 (1.45)	
Buy Side	1.16 (1.51)	1.43 (1.87)	1.40 (1.83)	1.24 (1.59)

Table 5.3 Annualized percentage turnover of the *Cross-Sectional Quantile* and *Single-Stock Quantile* portfolios.

	SS	*S*	*H*	*B*	*SB*	*Sell-Side*	*Buy-Side*
CSQ	136.51	158.20	245.24	148.15	142.86	102.38	98.016
SSQ	164.51	236.37	149.70	241.56	167.48	200.81	136.16

as *B*; if $2 < \overline{A}_t^i \leq 2.5$ as *H*; if $2.5 < \overline{A}_t^i \leq 3$ as *S*; and a *SS* whenever $\overline{A}_t^i > 3$. The downward shift accounts for the observed over optimistic recommendation scores provided by experts (see Barber et al. (2001) and the references therein). Using the expert recommendation system, we build six-month rebalanced portfolios exactly as we did for our systems and perform the regression in (5.1). The results in Table 5.1 show that the experts' *Buy Side* does not provide significant abnormal returns. In fact, abnormal returns are observed for the experts' *Sell Side* portfolio. Moreover, if the Sharpe and Sortino ratio of the *Sell Side*, *Hold* and *Buy Side* are considered, it turns out to be more clear, see Table 5.2, how Analysts does not provide a consistent recommendation system, i.e., a system where the more profitable portfolios are those with higher recommendations.

The previous calculated returns are the gross of all trading costs. To assess the size of these costs, we calculate a measure of annual turnover. Let t_h with $h = 1, \ldots, 19$ be the rebalancing dates, N_h^p be the number of companies in portfolio p at date t_h and δ_{i,t_h}^p be equal to 1 if company i is in portfolio p at date t_h and zero otherwise. Turnover at date t_{h+1} is calculated as

$$\mathrm{TO}_{t_{h+1}}^p := \sum_i \left| \frac{\delta_{i,t_h}^p}{N_h^p} - \frac{\delta_{i,t_{h+1}}^p}{N_{h+1}^p} \right|$$

where the sum is for all companies composing our universe. Annualized total turnover TO^p is twice the average of the previous quantity across the entire period. The values are reported in Table 5.3.

Table 5.4 The annualized abnormal returns and adjusted annualized percentage Sharpe ratios for the *Buy Side* portfolios as a function of $\text{RTC}^p(\%)$ for both *CSQ* and *SSQ* systems. The Sharpe ratios that are significantly different from that of *Our Universe*, according to the Ledoit and Wolf (2008) and Ardia and Boudt (2018) tests, at 10%, 5%, 1% and 0.1% level are marked with '$\circ$', '$*$', '$**$' and '$***$' respectively.

	CSQ		*SSQ*	
$\text{RTC}^p(\%)$	Ann. Abn. Return (%)	Ann. Sharpe Ratio (%)	Ann. Abn. Return (%)	Ann. Sharpe Ratio (%)
0.00	6.10	1.43 **	5.84	1.40 *
0.31	5.79	1.41 **	5.42	1.37 $\circ$
0.63	5.49	1.39 *	4.99	1.34
0.94	5.18	1.37 *	4.57	1.31
1.25	4.87	1.35 $\circ$	4.14	1.28
1.31	4.81	1.34	4.06	1.28
1.38	4.75	1.34	3.97	1.27
1.56	4.57	1.32	3.72	1.25
1.88	4.26	1.30	3.29	1.22
2.19	3.95	1.28	2.87	1.20
2.50	3.65	1.26	2.44	1.17
2.81	3.34	1.24	2.01	1.14

The transaction cost of portfolio p is computed as the product of the annualized turnover and the round-trip cost (RTC) (see e.g., Baule and Wilke, 2016, and references therein) and the "critical" round-trip cost, RTC^p_{crit}, which is the rebalancing cost that makes the net abnormal return of the portfolio equal to that of the *Our universe* benchmark:

$$\alpha^p - \text{RTC}^p_{crit} \cdot \text{TO}^p = \alpha^{\text{Our universe}}.$$

The critical round-trip cost is equal to 2.81% for the buy side *CSQ* and 1.84% for the buy side *SSQ* portfolios, the only ones with an abnormal return above that of the benchmark. Table 5.4 displays the adjusted (for transaction costs) annualized abnormal returns and the adjusted Sharpe ratio (i.e., computed using the adjusted returns, defined as the difference between the actual returns and transaction costs) for different levels of round-trip costs. As can be seen, the considered portfolios remain profitable even when transaction costs are fairly high. The buy side *CSQ*, with the lowest turnover, is less sensitive to transaction costs.

5.2.1 Building z-scores from the analysts' recommendations

How are affected the results in the previous analysis if we replace in the definition of the z-score the mean and the standard deviation of the fair-value distribution with the average and the standard deviation of the price targets given by the I/B/E/S database?

Table 5.5 This table reports the Sharpe and the Sortino ratios for the H, *Sell Side* and *Buy Side* portfolios constructed according to *Cross-Sectional-IBES*, *Cross-Sectional Quantile* and *Single-Stock Quantile*. Sharpe and the Sortino ratios for *Our universe* is also reported.

Sharpe-ratio (Sortino-ratio)	$CS\text{-}IBES$	CSQ	SSQ	*Our universe*
Sell Side	1.19 (1.53)	1.05 (1.36)	0.97 (1.28)	
H	1.10 (1.43)	1.15 (1.48)	1.11 (1.45)	
Buy Side	1.30 (1.70)	1.43 (1.87)	1.40 (1.83)	1.24 (1.59)

We denote this recommendation approach *Cross-Sectional-IBES* ($CS\text{-}IBES$)[5]. Table 5.5 summarizes the results. Taking *Our-universe* as baseline, the buy-side of CSQ and SSQ present the higher Sharpe Ratio and Sortino gains. We emphasize that also $CS\text{-}IBES$ provides coherent results for *Buy Side* and *Sell Side* portfolios, in contrast with the methodology employed by Barber et al. (2001). Remarkably, the difference between the Sharpe Ratio of the $CS\text{-}IBES$ and CSQ buy-side is statistically significant at the 5% level, whereas it is not significant the difference between the $CS\text{-}IBES$ and *Our-universe* buy-side, again according to the test of Ledoit and Wolf (2008) and Ardia and Boudt (2018). The investigation carried out in this section shows that we can improve the research design of Barber et al. (2001) using the information provided by the average and standard deviation of I/B/E/S's price targets with the methodology of the cross-sectional z-score recommendations. We have shown how to use the uncertainty provided by analysts and thus how to improve their recommendations, by relying on the CSQ methodology. In particular, this exhibits the robustness of the cross-sectional approach to issue consistent recommendation system.

We have concluded the first part of the thesis by proposing two recommendation systems based on the comparison of observed market prices with the fair value distributions derived by the SDCF approach presented in Chapter 3. For each recommendation system, we build buy and sell side portfolios and estimate the abnormal returns, both gross and net trading costs, earned from diverse investment strategies. Over all, the analyses show how the above proposed recommendations systems are able to provide profitable portfolio strategies. However, the execution and liquidation of several positions at the same time, e.g., the rebalancing of a portfolio which ensures a minimum exposition and high protection against high trading costs and arbitrageur, may involve sophisticated trading techniques. Therefore, the last part of the thesis is dedicated to the study of

[5]It is worth noting that one of the nice features of this method is that it cannot happen that for some rebalancing date portfolios with the most/least favorably recommended companies are empty, as it may be the case with the rescaling proposed by Barber et al. (2001).

optimal execution problems of asset portfolios in a market with many competitors. In particular, we show how to exploit the market impact due the agents trading activities to derive optimal liquidation strategies in the so called *market impact games* framework. Part of the following chapters is also dedicated to the analysis of some trading conditions which lead to instability, see Chapter 9.

Part II

Market Stability in Market Impact Games

We now consider the general problem of a set of agents trading a portfolio of assets in the presence of transient price impact and additional quadratic transaction costs and we study, with analytical and numerical methods, the resulting Nash equilibria. Extending significantly the framework of Schied and Zhang (2018) and Luo and Schied (2020), who considered the one asset case, we focus our attention on the conditions on the value of transaction cost making the trading profile of the agents, and as a consequence the price trajectory, wildly oscillating and the market unstable. We prove the existence and uniqueness of the corresponding Nash equilibria for the related mean-variance optimization problem. We find that the presence of more assets and a large number of agents make the market more prone to large oscillations and instability. When the number of assets is fixed, a more complex structure of the cross-impact matrix, i.e. the existence of multiple factors for liquidity, makes the market less stable compared to the case when a single liquidity factor exists.

CHAPTER 6

Market Impact Games

Let us consider two traders who want to trade simultaneously a certain number of shares, minimizing the trading costs. Since the trading of one agent affects the price, the other agent must take into account the presence of the former in optimizing her execution. This problem is termed *market impact game* and has received considerable attention in recent years (Carlin et al. (2007),Schöneborn (2008),Moallemi et al. (2012), Lachapelle et al. (2016),Strehle (2017a,b), Schied and Zhang (2018)). The seminal paper by Schied and Zhang, (Schied and Zhang (2018)), considers a market impact game between two identical agents trading the same asset in a given time period.

6.1 The Schied and Zhang framework

When none of the two agents trade, the price dynamics is described by the so called unaffected price process S_t^0 which is a right-continuous martingale defined on a given probability space $(\Omega, (\mathscr{F}_t)_{t\geq 0}, \mathscr{F}, \mathbb{P})$. A trader wants to unwind a given initial position with inventory Z, where a positive (negative) inventory means a short (long) position, during a given trading time grid $\mathbb{T} = \{t_0, t_1, \ldots, t_N\}$, where $0 = t_0 < t_1 < \cdots < t_N = T$ and following an admissible strategy, which is defined as follows:

Definition 6.1 (Admissible Strategy)**.** Given $\mathbb{T}$ and Z, an *admissible trading strategy* for $\mathbb{T}$ and $Z \in \mathbb{R}$ is a vector $\boldsymbol{\zeta} = (\zeta_0, \zeta_1, \ldots, \zeta_N)$ of random variables such that:

- $\zeta_k \in \mathscr{F}_{t_k}$ and bounded, $\forall k = 0, 1, \ldots, N$.
- $\zeta_0 + \zeta_1 + \cdots + \zeta_N = Z$.

The random variable ζ_k represents the order flow at trading time t_k where positive (negative) flow corresponds to a sell (buy) trade of volume $|\zeta_k|$. We denote with X_1 and X_2 the initial inventories of the two considered agents playing the game and with $\Xi = (\xi_{i,k}) \in \mathbb{R}^{2\times(N+1)}$ the matrix of the respective strategies, where $\boldsymbol{\xi}_{1,\cdot} = \{\xi_{1,k}\}_{k\in\mathbb{T}}$ and $\boldsymbol{\xi}_{2,\cdot} = \{\xi_{2,k}\}_{k\in\mathbb{T}}$ are the strategies of trader 1 and 2, respectively. Traders are subject to fees and transaction costs and their objective is to minimize them by optimizing the execution. As customary in the literature, the costs are modeled by two components. The first one is a temporary impact component modeled by a quadratic term $\theta\xi_{j,k}^2$,

respectively for trader j, which does not affect the price dynamics. This is sometimes called slippage and depends on the immediate liquidity present in the order book. Notice that, as discussed in Schied and Zhang (2018), this term can also be interpreted as a quadratic transaction fee. Here we do not specify exactly what this term represents, sticking to the mathematical modeling approach of Schied and Zhang.

The second component is related to permanent impact and affects future price dynamics. Following Schied and Zhang (2018), we consider the celebrated transient impact model of Bouchaud et al. (2004, 2009), which describes the price process S_t^{Ξ} affected by the strategies Ξ of the two traders, i.e.,

$$S_t^{\Xi} = S_t^0 - \sum_{t_k<t} G(t-t_k)(\xi_{1,k}+\xi_{2,k}), \quad \forall\, t \in \mathbb{T},$$

where $G : \mathbb{R}_+ \to \mathbb{R}_+$ is the so called *decay kernel*, which describes the lagged price impact of a unit buy or sell order over time. Usual assumptions on G are satisfied, i.e., it is convex, nonincreasing, nonconstant so that $t \mapsto G(|t|)$ is strictly positive definite in the sense of Bochner, see Alfonsi et al. (2012) and Schied and Zhang (2018). Notice that by choosing a constant kernel G, one recovers the celebrated Almgren-Chriss model (Almgren and Chriss (2001)).

The cost faced by each agent is the sum of the two components above. Specifically, let us denote with $\mathscr{X}(X,\mathbb{T})$ the set of admissible strategies for the initial inventory X on a specified time grid $\mathbb{T}$, the cost functions are defined following Schied and Zhang (2018):

Definition 6.2. Given $\mathbb{T} = \{t_0, t_1, \dots, t_N\}$, X_1 and X_2. Let $(\varepsilon_i)_{i=0,1,\dots N}$ be an i.i.d. sequence of Bernoulli $\left(\frac{1}{2}\right)$-distributed random variables that are independent of $\sigma(\bigcup_{t\geq 0}\mathscr{F}_t)$. Then the *cost of* $\boldsymbol{\xi}_{1,\cdot} \in \mathscr{X}(X_1,\mathbb{T})$ *given* $\boldsymbol{\xi}_{2,\cdot} \in \mathscr{X}(X_2,\mathbb{T})$ is defined as

$$C_{\mathbb{T}}(\boldsymbol{\xi}_{1,\cdot}|\boldsymbol{\xi}_{2,\cdot}) = \sum_{k=0}^{N}\left(\frac{G(0)}{2}\xi_{1,k}^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\xi_{1,k} + \varepsilon_k G(0)\xi_{1,k}\xi_{2,k} + \theta\xi_{1,k}^2\right) + X_1 S_0^0$$

and the *costs of* $\boldsymbol{\xi}_{2,\cdot}$ *given* $\boldsymbol{\xi}_{1,\cdot}$ are

$$C_{\mathbb{T}}(\boldsymbol{\xi}_{2,\cdot}|\boldsymbol{\xi}_{1,\cdot}) = \sum_{k=0}^{N}\left(\frac{G(0)}{2}\xi_{2,k}^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\xi_{2,k} + (1-\varepsilon_k)G(0)\xi_{1,k}\xi_{2,k} + \theta\xi_{2,k}^2\right) + X_2 S_0^0.$$

Thus the execution priority at time t_k is given to the agent who wins an independent coin toss game, represented by a Bernoulli variable ε_k, which is a fair game in the framework of Schied and Zhang (2018).

Given the time grid $\mathbb{T} = \{t_0, t_1, \dots, t_N\}$ and the initial values $X_1, X_2 \in \mathbb{R}$, we define the *Nash Equilibrium* as a pair $(\boldsymbol{\xi}_{1,\cdot}^*, \boldsymbol{\xi}_{2,\cdot}^*)$ of strategies in $\mathscr{X}(X_1,\mathbb{T}) \times \mathscr{X}(X_2,\mathbb{T})$ such

that

$$\mathbb{E}[C_{\mathbb{T}}(\boldsymbol{\xi}^*_{1,\cdot}|\boldsymbol{\xi}^*_{2,\cdot})] = \min_{\boldsymbol{\xi}_{1,\cdot}\in\mathscr{X}(X_1,\mathbb{T})} \mathbb{E}[C_{\mathbb{T}}(\boldsymbol{\xi}_{1,\cdot}|\boldsymbol{\xi}^*_{2,\cdot})] \text{ and}$$
$$\mathbb{E}[C_{\mathbb{T}}(\boldsymbol{\xi}^*_{2,\cdot}|\boldsymbol{\xi}^*_{1,\cdot})] = \min_{\boldsymbol{\xi}_{2,\cdot}\in\mathscr{X}(X_2,\mathbb{T})} \mathbb{E}[C_{\mathbb{T}}(\boldsymbol{\xi}_{2,\cdot}|\boldsymbol{\xi}^*_{1,\cdot})].$$

One of main results of Schied and Zhang (2018) is the proof, under general assumptions, of the existence and uniqueness of the Nash equilibrium. Moreover, they showed that this equilibrium is deterministically given by a linear combination of two constant vectors, namely fundamental solutions $\boldsymbol{v}$ and $\boldsymbol{w}$ which are defined as

$$\boldsymbol{v} = \frac{1}{\boldsymbol{e}^T(\Gamma_\theta + \widetilde{\Gamma})^{-1}\boldsymbol{e}}(\Gamma_\theta + \widetilde{\Gamma})^{-1}\boldsymbol{e}$$
$$\boldsymbol{w} = \frac{1}{\boldsymbol{e}^T(\Gamma_\theta - \widetilde{\Gamma})^{-1}\boldsymbol{e}}(\Gamma_\theta - \widetilde{\Gamma})^{-1}\boldsymbol{e}.$$

and $\boldsymbol{e} = (1,\dots,1)^T \in \mathbb{R}^{N+1}$. The kernel matrix $\Gamma \in \mathbb{R}^{(N+1)\times(N+1)}$ is given by

$$\Gamma_{ij} = G(|t_{i-1} - t_{j-1}|), \quad i,j = 1,2,\dots,N+1,$$

and for $\theta \geq 0$ it is $\Gamma_\theta := \Gamma + 2\theta I$, and the matrix $\widetilde{\Gamma}$ is given by

$$\widetilde{\Gamma}_{ij} = \begin{cases} \Gamma_{ij} & \text{if } i > j \\ \frac{1}{2}G(0) & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

As showed by Schied and Zhang (2018) all these matrices are definite positive.

Theorem 6.3 (Nash Equilibrium for Market Impact Games, Schied and Zhang (2018))**.** *For any strictly positive definite decay kernel G, time grid $\mathbb{T}$, parameter $\theta \geq 0$, and initial values $X_1, X_2 \in \mathbb{R}$, there exists a unique Nash equilibrium $(\boldsymbol{\xi}^*, \boldsymbol{\eta}^*) \in \mathscr{X}(X_1,\mathbb{T}) \times \mathscr{X}(X_2,\mathbb{T})$. The optimal strategies $\boldsymbol{\xi}^*$ and $\boldsymbol{\eta}^*$ are deterministic and given by*

$$\boldsymbol{\xi}^* = \frac{1}{2}(X_1 + X_2)\boldsymbol{v} + \frac{1}{2}(X_1 - X_2)\boldsymbol{w} \tag{6.1}$$
$$\boldsymbol{\eta}^* = \frac{1}{2}(X_1 + X_2)\boldsymbol{v} - \frac{1}{2}(X_1 - X_2)\boldsymbol{w}, \tag{6.2}$$

An interesting result of Schied and Zhang (2018) concerns the stability of the Nash equilibrium related to the transaction costs parameter θ and the decay kernel G. Generically, following Schied and Zhang (2018), we say that a market is *unstable* if the trading strategies at the Nash equilibrium exhibit spurious oscillations, i.e., if there exists a sequence of trading times such that the orders are consecutively composed by buy and

sell trades, for all initial inventories X_1 and Y_2. In the optimal execution literature such behavior is termed *transaction triggered price manipulation*, see Alfonsi et al. (2012). Figure 6.1 shows the simulation of the price process under the Schied and Zhang model when both investors have an inventory equal to 1 for two values of θ. The unaffected price process is a simple random walk with zero drift and constant volatility and the trading of the two agents, according to the Nash equilibrium, modifies the price path. For small θ (top panel) the affected price process exhibits wild oscillations, while when θ is large (bottom panel) the irregular behavior disappears[1].

Thus, Schied and Zhang (2018) showed, when the trading time grid is equispaced, $\mathbb{T}_N$, and under general assumptions on G, the existence of a critical value $\theta^* = G(0)/4$ such that for $\theta < \theta^*$ the equilibrium strategies exhibit oscillations of buy and sell orders for both traders. Hence, the behavior at zero of the kernel function plays a relevant role for the equilibrium stability.

Proposition 6.4 (Instability in Market Impact Games, Schied and Zhang (2018))**.** *Suppose that G is a continuous, positive definite, strictly positive, log-convex decay kernel and that $\mathbb{T}_N = \{\frac{k}{N} | k = 0, 1, \dots, N\}$ denotes the equidistant time grid. Then, the following conditions are equivalent:*

i) For every $N \in \mathbb{N}$ and $T > 0$, all components of $\boldsymbol{w}$ are nonnegative.

ii) $\theta \geq \theta^* = G(0)/4$.

If, moreover, G is an exponential decay kernel, i.e., $G(t) = \lambda e^{-\rho t} + \gamma$ for $\lambda, \rho > 0$ and $\gamma \geq 0$, then conditions i) and ii) are equivalent to:

iii) For every $N \in \mathbb{N}$ and $T > 0$, all components of $\boldsymbol{v}$ are nonnegative.

As mentioned in the introduction, this result has been proved for a market with only $M = 1$ asset, two ($J = 2$) risk-neutral traders. Now, we recall the extension of this framework in a multi-agent market ($J > 2$) of Luo and Schied (2020). Then, we first extend their framework in the multi-asset ($M > 1$) case, where we show the existence and uniqueness of the related Nash equilibrium, and finally we generalize the stability result of Schied and Zhang (2018) in the multi-asset case.

6.2 The Luo and Schied multi-agent market impact model

The Luo and Schied (2020) model is an extension of the Schied and Zhang (2018) model where J risk-averse traders want to trade the same asset. The unaffected price

[1]Moreover, we observe that the presence of spurious oscillations in the price dynamics may affect the consistency of the spot volatility estimation. Indeed, these oscillations act as a market microstructure noise, even if this noise is caused by the oscillations of a deterministic trend, while usually it is characterized by some additive noise term. In particular, we find that when θ is close to zero the noise is amplified by spurious oscillations, while for sufficiently large θ these oscillations do not compromise the consistency of the spot volatility.

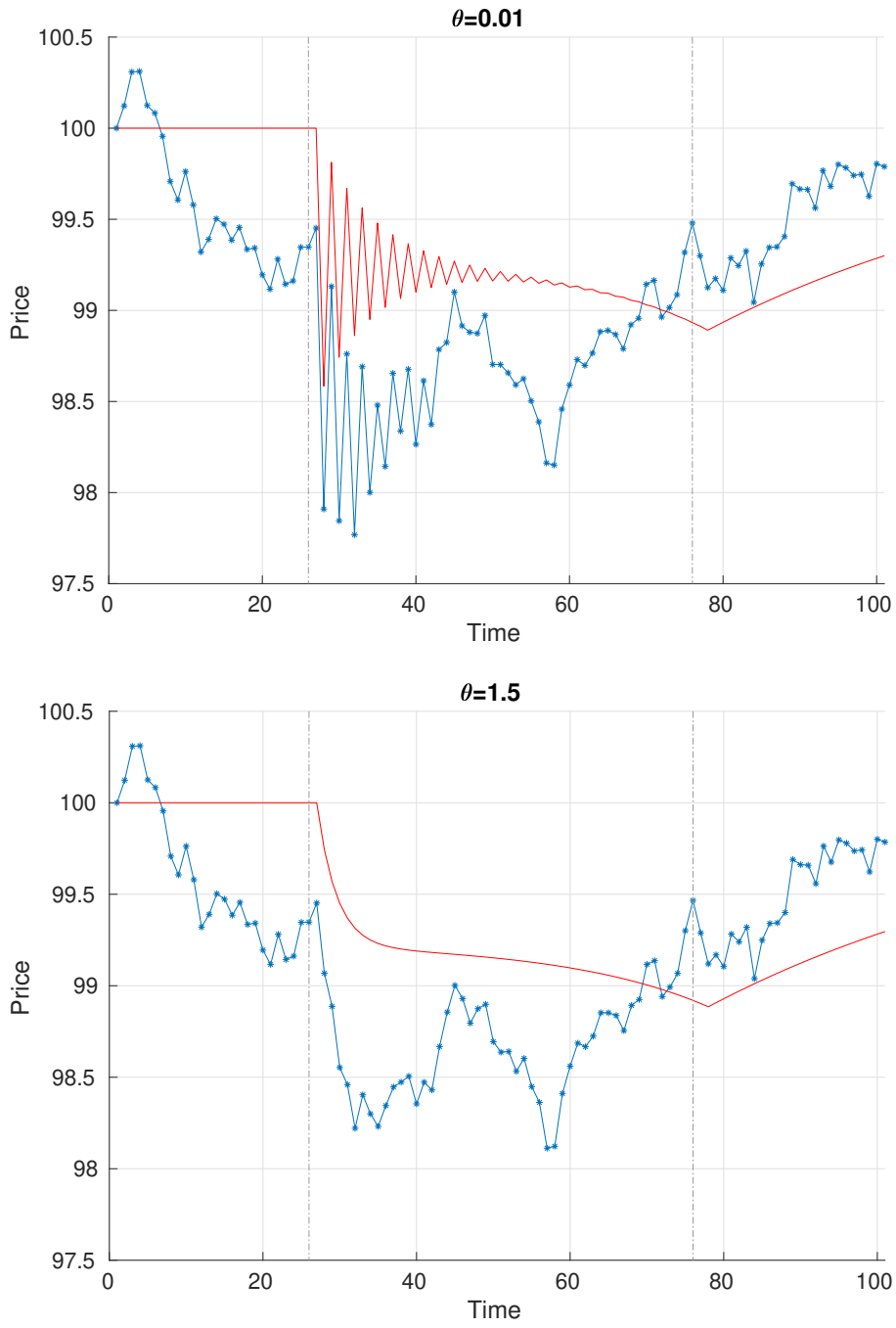


Fig. 6.1 Blue lines exhibit the price process when both agents have inventory equals to 1. The top (bottom) panel shows the dynamics when $\theta = 0.01$ ($\theta = 1.5$). The trading time grid has $N + 1 = 51$ points, $G(t) = \exp(-t)$, the volatility of the unaffected price process is fixed to 1 and $S_0 = 100$. The vertical grey dotted lines delineates the trading session. The red lines shows the drift dynamics due to trading.

process S^0_t is always assumed to be a right continuous martingale in a suitable filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ and it is also required that $S^0_\cdot$ is a square-integrable process. As before, let $\mathbb{T} = \{t_0, t_1, \ldots, t_N\}$ be the trading time grid. Consistently with the previous notation, we denote with $\Xi = (\xi_{j,k}) \in \mathbb{R}^{J\times(N+1)}$ the matrix of all strategies,

where $\xi_{j,k}$ is the order flow of agent j at time t_k, so that the affected price process is defined as

$$S_t^{\Xi} := S_t^0 - \sum_{t_k<t} G(t-t_k) \cdot \sum_{j=1}^{J} \xi_{j,k},$$

where G is the decay kernel. The generalization of admissible strategy is straightforward, indeed if X_j denotes the inventory of the j-th agent, Ξ is admissible for $\boldsymbol{X} \in \mathbb{R}^J$ and $\mathbb{T}$, if $\boldsymbol{\xi}_{j,\cdot}$ is admissible for X_j and $\mathbb{T}$ for each j according to definition 6.1, i.e., it is adapted to the filtration, bounded and $\sum_{k=0}^{N} \xi_{j,k} = X_j$. The set of admissible strategy is denoted as $\mathscr{X}(\boldsymbol{X}, \mathbb{T})$. Then, if we consider all the possible time priorities among the J traders at each time step, i.e., all the possible permutations that determine the time priority for each trading time t_k which we may assume to be equiprobable, it is possible to generalize the previous definition of liquidation cost for a trader strategy, see Luo and Schied (2020) for further details. We denote $\Xi_{-j,\cdot}$ the matrix Ξ where the j-th row is eliminated.

Definition 6.5 (Luo and Schied (2020))**.** Given a time grid $\mathbb{T}$, the execution costs of a strategy $\boldsymbol{\xi}_{j,\cdot}$ given all other strategies $\boldsymbol{\xi}_{l,\cdot}$ where $l \neq j$ is defined as

$$C_{\mathbb{T}}(\boldsymbol{\xi}_{j,\cdot}|\Xi_{-j,\cdot}) = \sum_{k=0}^{N} \left(\frac{G(0)}{2} \xi_{j,k}^2 - S_{t_k}^{\Xi} \xi_{j,k} + \frac{G(0)}{2} \sum_{l \neq j} \xi_{j,k} \xi_{l,k} + \theta\ \xi_{j,k}^2 \right),$$

where $\theta \geq 0$.

In the framework of Schied and Zhang (2018) we have two risk-neutral agents which want to minimize the expected costs of a strategy, i.e., implementation shortfall orders. Now, following Luo and Schied (2020), we consider the agents' risk aversion by introducing the mean-variance and expected utility functionals, respectively

$$MV_\gamma(\boldsymbol{\xi}_{j,\cdot}|\Xi_{-j,\cdot}) := \mathbb{E}[C_{\mathbb{T}}(\boldsymbol{\xi}_{j,\cdot}|\Xi_{-j,\cdot})] + \frac{\gamma}{2}\mathrm{Var}[C_{\mathbb{T}}(\boldsymbol{\xi}_{j,\cdot}|\Xi_{-j,\cdot})], \tag{6.3}$$

$$U_\gamma(\boldsymbol{\xi}_{j,\cdot}|\Xi_{-j,\cdot}) := \mathbb{E}[u_\gamma(-C_{\mathbb{T}}(\boldsymbol{\xi}_{j,\cdot}|\Xi_{-j,\cdot}))], \tag{6.4}$$

where γ is the risk-aversion parameter and $u_\gamma(x)$ is the CARA utility function,

$$u_\gamma(x) = \begin{cases} \frac{1}{\gamma}(1 - e^{-\gamma x}) & \text{if } \gamma > 0, \\ -x & \text{if } \gamma = 0. \end{cases}$$

As usual, see e.g. Almgren and Chriss (2001), the minimization of the mean-variance functional is restricted to deterministic admissible strategies, which is denoted as $\mathscr{X}_{\mathrm{det}}(\boldsymbol{X}, \mathbb{T})$. All agents are assumed to have the same risk-aversion $\gamma \geq 0$, see Luo and Schied (2020) for further details. Moreover, they introduced the corresponding Nash equilibrium for the previously defined functionals.

Definition 6.6 (Luo and Schied (2020))**.** Given the time grid $\mathbb{T}$ and initial inventories $\boldsymbol{X} \in \mathbb{R}^J$ for J traders with risk aversion parameter $\gamma \neq 0$, then:

- a *Nash Equilibrium for mean-variance optimization* is a matrix of strategies $\Xi^* \in \mathscr{X}_{\text{det}}(\boldsymbol{X}, \mathbb{T})$ such that each row $\boldsymbol{\xi}^*_{j,\cdot}$ minimizes the mean-variance functional $MV_\gamma(\boldsymbol{\xi}_{j,\cdot}|\Xi^*_{-j,\cdot})$ over $\boldsymbol{\xi}_{j,\cdot} \in \mathscr{X}_{\text{det}}(X_j, \mathbb{T})$;
- a *Nash Equilibrium for CARA expected utility maximization* is a matrix of strategies $\Xi^* \in \mathscr{X}(\boldsymbol{X}, \mathbb{T})$ such that each row $\boldsymbol{\xi}^*_{j,\cdot}$ maximizes the CARA expected utility functional $U_\gamma(\boldsymbol{\xi}_{j,\cdot}|\Xi^*_{-j,\cdot})$ over $\boldsymbol{\xi}_{j,\cdot} \in \mathscr{X}(X_j, \mathbb{T})$.

In particular, Luo and Schied (2020) have shown that when the decay kernel is strictly positive definite and for any $\mathbb{T}$, parameters $\theta, \gamma \geq 0$ and initial inventories $\boldsymbol{X} \in \mathbb{R}^J$, there exists a unique Nash equilibrium for the mean-variance optimization. In particular, let $\boldsymbol{v}$, $\boldsymbol{w}$ be the fundamental solutions defined as

$$\begin{aligned}\boldsymbol{v} &= \frac{1}{\boldsymbol{e}^T[\Gamma^{\gamma,\theta} + (J-1)\widetilde{\Gamma}]^{-1}\boldsymbol{e}}[\Gamma^{\gamma,\theta} + (J-1)\widetilde{\Gamma}]^{-1}\boldsymbol{e} \\ \boldsymbol{w} &= \frac{1}{\boldsymbol{e}^T[\Gamma^{\gamma,\theta} - \widetilde{\Gamma}]^{-1}\boldsymbol{e}}[\Gamma^{\gamma,\theta} - \widetilde{\Gamma}]^{-1}\boldsymbol{e},\end{aligned}$$

where, if $\varphi(t) := \text{Var}(S^0_t)$, for $t \geq 0$, the matrix $\Gamma^{\gamma,\theta}$ is defined for $\theta, \gamma \geq 0$ as

$$\Gamma^{\gamma,\theta}_{i,j} := (\Gamma_\theta)_{i,j} + \gamma\varphi(t_{i-1} \wedge t_{j-1}), \quad i,j = 1,2,\ldots,N+1,$$

where Γ_θ is the previously defined kernel matrix. We recall that S_0 is a Bachelier model, if $S^0_t = S_0 + \sigma B_t$, for $t \geq 0$, where $S_0, \sigma > 0$ are constants and B_t is a standard Brownian motion, Then, the following result holds, see Luo and Schied (2020) for further details.

Theorem 6.7 (Nash Equilibrium for Multi-Agents Market Impact Games, Luo and Schied (2020))**.** *If the decay kernel is strictly positive definite, then, for any time grid* $\mathbb{T}$*, parameters* $\theta, \gamma \geq 0$*, initial inventories* $X_1, X_2 \ldots, X_J \in \mathbb{R}$*, and* $\overline{X} = \frac{1}{J}\sum_{j=1}^J X_j$*, the strategies*

$$\boldsymbol{\xi}^*_{j,\cdot} = \overline{X}\boldsymbol{v} + (X_j - \overline{X})\boldsymbol{w}, \quad j = 1,2,\ldots,J, \tag{6.5}$$

form the unique Nash equilibrium for mean-variance optimization. If, moreover, S_0 *is a Bachelier model, then the strategies* (6.5) *also form a Nash equilibrium for CARA expected utility maximization and it is unique if we restrict all trader strategies to be deterministic.*

CHAPTER 7

Multi-Asset Market Impact Games

We now extend the previous framework allowing the J agents to trade a portfolio of $M > 1$ assets. Indeed, agents often liquidate portfolio positions, which accounts in trading simultaneously many assets. In general, the optimal execution of a portfolio is different from many individual asset optimal executions, because of (i) correlation in asset prices, (ii) commonality in liquidity across assets (Chordia et al. (2000)), and (iii) cross-impact effects. In the following we will focus mainly on the third effect, even if disentangling them is a challenging statistical problem and we will discuss its relations with the correlation in asset prices which ensure the existence of Nash equilibrium.

To proceed, we first extend the notion of admissible strategy to the multi-asset case. A strategy for J traders during the trading time interval $\mathbb{T}$ for M assets is a multidimensional array $\Xi = (\xi_{i,j,k}) \in \mathbb{R}^{M\times J\times(N+1)}$, where $\xi_{i,j,k}$ is the strategy for the j-th trader in the i-th asset at time step k. Straightforwardly, given a fixed time grid $\mathbb{T}$ and initial inventory $X \in \mathbb{R}^{M\times J}$, where each column j contains the inventories of trader j for the M assets, a strategy Ξ of random variables is admissible for X if i) for all time step k, $\Xi_{\cdot,\cdot,k}$ is $\mathscr{F}_{t_k}$-measurable and bounded and ii) $\sum_{k=0}^{N} \boldsymbol{\xi}_{\cdot,j,k} = \boldsymbol{X}_j \in \mathbb{R}^M$ for each j, where $\boldsymbol{X}_j$ is the j-th column of X.

The second important point is that the trading of one asset modifies also the price of the other asset(s). This effect is termed *cross-impact*. While *self-impact* may be attributed to a mechanical and induced consequence of the order book, the cross-impact may be understood as an effect related to mispricing in correlated assets which are exploited by arbitrageurs betting on a reversion to normality, see Almgren and Chriss (2001) and Schneider and Lillo (2019) for further details. Cross-impact has been empirically studied recently, see e.g. Mastromatteo et al. (2017), Schneider and Lillo (2019) and its role in optimal execution has been highlighted in Tsoukalas et al. (2019).

Mathematically cross-impact is modeled by introducing a function $\mathcal{Q} : \mathbb{R}_+ \times \mathbb{R}^M \to \mathbb{R}^M$ describing how the trading of the M assets affect their prices at a certain future time. Schneider and Lillo (2019) have discussed necessary conditions for the absence of price manipulation for multi-asset transient impact models. They have shown that the cross-impact function need to be symmetric and linear in order to avoid arbitrage and manipulations. Moreover, following example 3.1 of Alfonsi et al. (2016) and as empirically

observed by Mastromatteo et al. (2017), we assume the same temporal dependence of G among the assets. Then, we assume that $\mathcal{Q} = Q \cdot G(t)$ where Q is linear and symmetric, i.e., $Q \in \mathbb{R}^{M\times M}$ and $Q = Q^T$ and $G : \mathbb{R}_+ \to \mathbb{R}_+$. Also, we assume that Q is a nonsingular matrix. Therefore, the price process during order execution is defined as

$$\boldsymbol{S}_t^{\Xi} = \boldsymbol{S}_t^0 - \sum_{t_k<t} G(t-t_k)\cdot Q \cdot \sum_{j=1}^{J} \boldsymbol{\xi}_{\cdot,j,k}$$

where we refer to $Q \in \mathbb{R}^{M\times M}$ as the cross-impact matrix, $\boldsymbol{S}_t^0 \in \mathbb{R}^M$ is the unaffected price process which is assumed to be a right-continuous martingale defined on a suitable filtered probability space and it is a square-integrable process.

If for each asset the time priority among the traders is determined by considering all the possible permutations of agents for each trading time t_k, then, following the same motivation of Schied and Zhang (2018) and Luo and Schied (2020), the definition 6.5 of liquidation cost is generalized as follows:

Definition 7.1 (Execution Cost)**.** Given a time grid $\mathbb{T}$ and $\theta \geq 0$, the execution cost of a strategy $\Xi_{\cdot,j,\cdot}$ given all other strategies $\Xi_{\cdot,l,\cdot}$ where $l \neq j$ is defined as

$$C_{\mathbb{T}}(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot}) = \sum_{k=0}^{N}\Bigg(\frac{G(0)}{2}\langle Q\boldsymbol{\xi}_{\cdot,j,k},\boldsymbol{\xi}_{\cdot,j,k}\rangle - \langle \boldsymbol{S}_{t_k}^{\Xi},\boldsymbol{\xi}_{\cdot,j,k}\rangle + \\ + \frac{G(0)}{2}\sum_{l\neq j}\langle Q\boldsymbol{\xi}_{\cdot,l,k},\boldsymbol{\xi}_{\cdot,j,k}\rangle + \theta\ \langle \boldsymbol{\xi}_{\cdot,j,k},\boldsymbol{\xi}_{\cdot,j,k}\rangle_{\cdot}\Bigg).$$

The previous definition is motivated by the following argument. When only agent j trades, the prices are moved from $\boldsymbol{S}_{t_k}^{\Xi}$ to $\boldsymbol{S}_{t_k+}^{\Xi} = \boldsymbol{S}_{t_k}^{\Xi} - G(0)Q\boldsymbol{\xi}_{\cdot,j,k}$. However, the order is executed at the average price and the player incurs in the expenses

$$-\frac{1}{2}\langle(\boldsymbol{S}_{t_k}^{\Xi} + \boldsymbol{S}_{t_k+}^{\Xi}),\boldsymbol{\xi}_{\cdot,j,k}\rangle = \frac{G(0)}{2}\langle Q\boldsymbol{\xi}_{\cdot,j,k},\boldsymbol{\xi}_{\cdot,j,k}\rangle - \langle \boldsymbol{S}_{t_k}^{\Xi},\boldsymbol{\xi}_{\cdot,j,k}\rangle.$$

Then, suppose that immediately after j the agent l place an order and the prices are moved linearly from $\boldsymbol{S}_{t_k+}^{\Xi}$ to $\boldsymbol{S}_{t_k+}^{\Xi} - G(0)Q\boldsymbol{\xi}_{\cdot,l,k}$, so the cost for l is given by:

$$-\frac{1}{2}\langle(\boldsymbol{S}_{t_k+}^{\Xi} + \boldsymbol{S}_{t_k+}^{\Xi}) - G(0)Q\boldsymbol{\xi}_{\cdot,l,k},\boldsymbol{\xi}_{\cdot,l,k}\rangle = \frac{G(0)}{2}\langle Q\boldsymbol{\xi}_{\cdot,l,k},\boldsymbol{\xi}_{\cdot,l,k}\rangle - \langle \boldsymbol{S}_{t_k}^{\Xi},\boldsymbol{\xi}_{\cdot,l,k}\rangle \\ + G(0)\langle Q\boldsymbol{\xi}_{\cdot,j,k},\boldsymbol{\xi}_{\cdot,l,k}\rangle.$$

The term $G(0)\langle Q\boldsymbol{\xi}_{\cdot,j,k},\boldsymbol{\xi}_{\cdot,l,k}\rangle$ is the additional cost due to the latency, where on average for each asset half of the times the order of agent j will be executed before the one of agent l, so that the latency costs for agent j at time step k is given by $\frac{G(0)}{2}\sum_{l\neq j}\langle Q\boldsymbol{\xi}_{\cdot,l,k},\boldsymbol{\xi}_{\cdot,j,k}\rangle$, see Luo and Schied (2020) for further details.

The mean-variance and CARA expected utility functionals are straightforwardly generalized using the previous defined execution cost. Indeed,

$$MV_\gamma(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot}) := \mathbb{E}[C_{\mathbb{T}}(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot})] + \frac{\gamma}{2}\mathrm{Var}[C_{\mathbb{T}}(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot})], \tag{7.1}$$

$$U_\gamma(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot}) := \mathbb{E}[u_\gamma(-C_{\mathbb{T}}(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot}))]. \tag{7.2}$$

Therefore, we may define the related Nash equilibria definitions:

Definition 7.2. Given the time grid $\mathbb{T}$ and initial inventories $X \in \mathbb{R}^{M\times J}$ for M assets and J traders with risk aversion parameter $\gamma \geq 0$, then:

- a *Nash Equilibrium for mean-variance optimization* is a multidimensional array of strategies $\Xi^* \in \mathscr{X}_{\mathrm{det}}(X, \mathbb{T})$ such that $\Xi^*_{\cdot,j,\cdot}$ minimizes the mean-variance functional $MV_\gamma(\Xi_{\cdot,j,\cdot}|\Xi^*_{\cdot,-j,\cdot})$ over $\Xi_{\cdot,j,\cdot} \in \mathscr{X}_{\mathrm{det}}(\boldsymbol{X}_j, \mathbb{T})$;

- a *Nash Equilibrium for CARA expected utility maximization* is a multidimensional array of strategies $\Xi^* \in \mathscr{X}(X, \mathbb{T})$ such that each row $\Xi^*_{\cdot,j,\cdot}$ maximizes the CARA expected utility functional $U_\gamma(\Xi_{\cdot,j,\cdot}|\Xi^*_{\cdot,-j,\cdot})$ over $\Xi_{\cdot,j,\cdot} \in \mathscr{X}(\boldsymbol{X}_j, \mathbb{T})$.

We recall that $\boldsymbol{S}^0_t$ follows a Bachelier model if $\boldsymbol{S}^0_t = \boldsymbol{S}_0 + L\boldsymbol{B}_t$ where $\boldsymbol{S}_0$ is a fixed vector and $\boldsymbol{B}_t$ is a multivariate (standard) Brownian motion, where its components are independent with unit variance so that the variance-covariance matrix of $\boldsymbol{S}^0_t$ is given by $\Sigma = LL^T$.

Remark 7.3. We are implicitly assuming that the strength of the impact of a single trader is independent from the number of agents simultaneously present. This is not necessarily true. For example, generalizing Kyle's model to the case when $J \geq 1$ symmetrically informed agents are simultaneously present, Bagnoli et al. (2001) shows that the Kyle's lambda, i.e. the proportionality factor between price impact and aggregated order flow, scales as $J^{-1/\alpha}$, where α is the exponent of the stable law describing the price and uninformed order flow distribution. Moreover if the second moment of both variables is finite, Bagnoli et al. (2001) shows that the Kyle's lambda scales as $1/\sqrt{J}$ (see also Lambert et al. (2018) for the non symmetrical case when distributions are Gaussian). In our impact model, this property can be modeled by assuming that the decay kernel depends on J as $G^{sc}(t) := J^{-\beta} \cdot G(t)$ where $G(t)$ is the standard non-scaled decay kernel and $\beta \geq 0$. The case $\beta = 0$ corresponds to the additive case, while for $\beta = 1$ the total instantaneous impact does not depend on the number of agents J. There are also some recent empirical evidences suggesting that the impact strength depends on the number of agents simultaneously trading. Figure 3 of Bucci et al. (2020) indicates that market

impact of a metaorder[1] decreases with the number of metaorders simultaneously present. In the following we first consider $G(t)$ independent from J, while we investigate in detail how the market stability is affected by the scaling parameter β in Section 9.2.2.

Remark 7.4. In the previous remark we have pointed out that there are some theoretical arguments suggesting that the kernel might depend and scale with the number of agents J. Therefore, we may expect an analogous argument for the cross-impact matrix and the number of assets M. In this merit, we may consider the multivariate Kyle model of Garcia del Molino et al. (2020), where we denote with Λ the cross-impact Kyle lambda matrix. In particular, Garcia del Molino et al. (2020) have shown the existence and uniqueness of the linear equilibrium of the multivariate Kyle model, where by Proposition 3.4 of Garcia del Molino et al. (2020), Λ satisfies $\frac{1}{4}\Sigma_0 = \Lambda\Omega\Lambda$ where Σ_0 and Ω are the covariance matrices of the fundamental price and of the bids of the noise trader, respectively. If we assume that these matrices can be decomposed as $\Sigma_0 = s_d I + s_n \boldsymbol{e}\boldsymbol{e}^T$ and $\Omega = \omega_d I + \omega_n \boldsymbol{e}\boldsymbol{e}^T$, where $\boldsymbol{e}$ is the vector with all components equal to one, then necessarily[2] $\Lambda = \lambda_d I + \lambda_n \boldsymbol{e}\boldsymbol{e}^T$. However, since $\frac{1}{4}\Sigma_0 = \Lambda\Omega\Lambda$, $\lambda_d = \sqrt{\frac{s_d}{4\omega_d}}$ which is independent from M (this was also empirically observed by Benzaquen et al. (2017)) and more interestingly, $\lim_{M\to\infty} \lambda_n = \lim_{M\to\infty}(-\lambda_d \pm \sqrt{\frac{s_d}{4\omega_n}})\frac{1}{M} = 0$. It is worth to nothing that even though $\omega_n = 0$, i.e., the order flow of the uninformed trader is uncorrelated, the cross-impact Kyle lambda, λ_n, is different from zero and $\lim_{M\to\infty} \lambda_n = \lim_{M\to\infty} \sqrt{\frac{s_n}{4\omega_d M}} = 0$.

Thus, the off-diagonal terms of the cross-impact matrix Q might scale as $1/M$. Therefore, asymptotically, when M becomes large, the cross-impact terms would vanish and Q would converge towards a diagonal matrix. As we will see later, this type of scaling may act as a stabilization effect. However, we will not further investigate this option in this thesis.

7.1 Nash equilibrium for the linear cross impact model

We now prove the existence and uniqueness of the Nash equilibrium in this multi-asset setting. This is achieved by using the spectral decomposition of Q to orthogonalize the assets, which we call "virtual" assets, so that the impact of the orthogonalized strategies on the virtual assets is fully characterized by the self-impact, i.e., the transformed cross impact matrix is diagonal. Thus, the existence and uniqueness of the Nash equilibrium derives immediately by following the same argument as in Schied and Zhang (2018) and Luo and Schied (2020).

[1] A metaorder is a sequence of trades executed in the same direction (either buys or sells) and originating from the same market participant. Thus in our framework each trader j executes a metaorder of size X_j.

[2] If $A = aI + b\boldsymbol{e}\boldsymbol{e}^T$ is a one-factor matrix, a particular case of rank-one update matrix, where $a, b \in \mathbb{R}$, then A^{-1} is one-factor and moreover L is also a one-factor matrix where $LL^T = A$. Therefore, by Theorem 3.5 equation (3.6) of Garcia del Molino et al. (2020) Λ is a one-factor matrix.

Remark 7.5. If we suppose that Q is the identity matrix, then the multi-asset market impact game is a straightforward generalization of the Luo and Schied (2020) model. Indeed, each order of the players for the i-th stock does not affect any other asset.

In general, if we assume that $\boldsymbol{S}_t^0$ has uncorrelated components, i.e., the variance-covariance matrix Σ is diagonal, then the following result holds.

Lemma 7.6 (Nash Equilibrium for Diagonal Cross-Impact Matrix)**.** *If $\boldsymbol{S}_t^0$ has uncorrelated components, for any strictly positive definite decay kernel G, time grid $\mathbb{T}$, parameters $\theta, \gamma \geq 0$, initial inventory $X \in \mathbb{R}^{M\times J}$ and diagonal positive cross impact matrix $D = diag(\lambda_1, \lambda_2, \ldots, \lambda_M)$, there exists a unique Nash Equilibrium $\Xi^* \in \mathscr{X}_{\text{det}}(X, \mathbb{T})$ for the mean-variance optimization problem and it is given by*

$$\boldsymbol{\xi}^*_{i,j,\cdot} = \overline{X}_{i,\cdot}\boldsymbol{v}_i + (X_{i,j} - \overline{X}_{i,\cdot})\boldsymbol{w}_i, \quad j = 1, 2, \ldots, J, \quad i = 1, 2, \ldots, M, \tag{7.3}$$

where $\overline{X}_{i,\cdot} = \frac{1}{J}\sum_{j=1}^{J} X_{i,j}$, $\boldsymbol{v}_i$ and $\boldsymbol{w}_i$ are the fundamental solutions associated with the decay kernel $G_i(t) = G(t) \cdot \lambda_i$ and same parameter θ. Moreover, if $\boldsymbol{S}_t^0$ follows a Bachelier model, then (7.3) *is also a Nash equilibrium for CARA expected utility maximization.*

Proof of Lemma 7.6. Since the cross-impact matrix is diagonal, each asset is not affected by the orders on other assets, i.e., the impact for each asset is provided only by the self-impact and there is no cross-impact effect. In particular,

$$C_{\mathbb{T}}(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot}) = \sum_{i=1}^{M} C_{\mathbb{T}}(\boldsymbol{\xi}_{i,j,\cdot}|\Xi_{i,-j,\cdot}; G_i),$$

where $C_{\mathbb{T}}(\boldsymbol{\xi}_{i,j,\cdot}|\Xi_{i,-j,\cdot}; G_i)$ is the liquidation cost of Definition 6.5 where the decay kernel is multiplied by λ_i. Moreover, the mean-variance functional can be splitted in the sum of mean-variance functionals of each asset i, i.e.,

$$MV_\gamma(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot}) = \sum_{i=1}^{M} MV_\gamma(\boldsymbol{\xi}_{i,j,\cdot}|\Xi_{i,-j,\cdot}; G_i),$$

where $MV_\gamma(\boldsymbol{\xi}_{i,j,\cdot}|\Xi_{i,-j,\cdot}; G_i)$ is the mean-variance functional defined in Eq. (6.3) with the related $C_{\mathbb{T}}(\boldsymbol{\xi}_{i,j,\cdot}|\Xi_{i,-j,\cdot}; G_i)$. Indeed, $MV_\gamma(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot}) = \mathbb{E}[C_{\mathbb{T}}(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot})] + \frac{\gamma}{2}\text{Var}[C_{\mathbb{T}}(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot})]$ and since $\mathbb{E}[\cdot]$ is a linear operator

$$E[C_{\mathbb{T}}(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot})] = \sum_{i=1}^{M} E[C_{\mathbb{T}}(\boldsymbol{\xi}_{i,j,\cdot}|\Xi_{i,-j,\cdot}; G_i)].$$

On the other hand, $\text{Var}[C_{\mathbb{T}}(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot})] = \text{Var}[\sum_{k=0}^{N}\langle \boldsymbol{S}^{\Xi}_{t_k}, \boldsymbol{\xi}_{\cdot,j,k}\rangle]$ because Ξ is deter-

ministic. Let us denote $Y_i = \sum_{k=0}^{N} S^{\Xi}_{t_k,i}\xi_{i,j,k}$, then

$$\mathrm{Var}\left[\sum_{k=0}^{N}\langle \boldsymbol{S}^{\Xi}_{t_k}, \boldsymbol{\xi}_{\cdot,j,k}\rangle\right] = \mathrm{Var}\left[\sum_{i=1}^{M} Y_i\right] = \sum_{i=1}^{M}\mathrm{Var}(Y_i) + \sum_{i\neq l}\mathrm{Cov}(Y_i, Y_l).$$

However, if $\mathrm{Cov}(Y_i, Y_l) = 0$ for $i \neq l$, then

$$\mathrm{Var}[C_{\mathbb{T}}(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot})] = \sum_{i=1}^{M}\mathrm{Var}[C_{\mathbb{T}}(\boldsymbol{\xi}_{i,j,\cdot}|\Xi_{i,-j,\cdot};G_i)],$$

where we used again that Ξ is deterministic. Therefore, the M multi-asset market impact game with J agents is equivalent to consider M stacked independent one-asset market impact game with J agents, where the decay kernel for each asset i is scaled by the corresponding diagonal element of D, λ_i, which preserves the strictly positive definite property since $\lambda_i > 0\ \forall i$. Thus, for each asset i and agent j the existence, uniqueness and the closed formula of Nash Equilibrium $\boldsymbol{\xi}^*_{i,j,\cdot}$ for the mean-variance optimization are straightforward from Theorem 2.4 of Luo and Schied (2020) where the decay kernel is multiplied by λ_i, respectively for each asset. Moreover, since $MV_\gamma(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot}) = \sum_{i=1}^{M} MV_\gamma(\boldsymbol{\xi}_{i,j,\cdot}|\Xi_{i,-j,\cdot};G_i)$ we may conclude. If $\boldsymbol{S}^0_\cdot$ follows a Bachelier model and Ξ is deterministic, then $C_{\mathbb{T}}(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot})$ is a Gaussian random variable, so that the mean-variance optimization and CARA expected utility maximization are equivalent over the class of deterministic strategies, indeed

$$U_\gamma(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot}) = u_\gamma(-MV_\gamma(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot})), \quad \gamma > 0,$$

$$\text{and } U_0(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot}) = -\mathbb{E}[C_{\mathbb{T}}(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot})], \quad \gamma = 0.$$

On the other hand, following the same reasoning of the proof of Theorem 2.4 of Luo and Schied (2020), when $\Xi_{\cdot,-j,\cdot}$ are deterministic, from Theorem 2.1 of Schied et al. (2010) if there exists a deterministic strategy $\Xi^*_{\cdot,j,\cdot}$ which maximizes the expected utility functional $U_\gamma(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot})$, over the class of deterministic strategies, then $\Xi^*_{\cdot,j,\cdot}$ is also a maximizer for the expected utility functional within the class of all adapted strategies. Then, we may use the same argument of Corollary 2.3 of Schied and Zhang (2017) to conclude that the Nash equilibrium for the mean-variance optimization problem form a Nash equilibrium for CARA expected utility maximization.

So, it remains to show that if $\boldsymbol{S}^0_\cdot$ has uncorrelated components, then $\mathrm{Cov}(Y_i, Y_l) = 0$ for $i \neq l$, where $Y_i = \sum_{k=0}^{N} S^{\Xi}_{t_k,i}\xi_{i,j,k}$. However, $S^{\Xi}_{t,i} = S^0_{t,i} - \sum_{t_k<t} G(t-t_k)\cdot\sum_{j=1}^{J}(Q\cdot$

$\boldsymbol{\xi}_{\cdot,j,k})_i$, where $(Q\cdot\boldsymbol{\xi}_{\cdot,j,k})_i$ denotes the i-th component of $Q\cdot\boldsymbol{\xi}_{\cdot,j,k}$, then

$$\begin{aligned} Y_i &= \sum_{k=0}^{N}\left[S^0_{t_k,i}\xi_{i,j,k} - \left(\sum_{t_k<t} G(t-t_k)\cdot\sum_{j=1}^{J}(Q\cdot\boldsymbol{\xi}_{\cdot,j,k})_i\right)\xi_{i,j,k}\right] \\ &= \sum_{k=0}^{N}\left[S^0_{t_k,i}\xi_{i,j,k}\right] - \sum_{k=0}^{N}\left[\left(\sum_{t_k<t} G(t-t_k)\cdot\sum_{j=1}^{J}(Q\cdot\boldsymbol{\xi}_{\cdot,j,k})_i\right)\xi_{i,j,k}\right] \end{aligned}$$

so since Ξ is deterministic and using the martingale property of $\boldsymbol{S}^0_\cdot$,

$$\begin{aligned} \mathrm{Cov}(Y_i,Y_l) &= \mathrm{Cov}\left(\sum_{k=0}^{N} S^0_{t_k,i}\xi_{i,j,k}, \sum_{h=0}^{N} S^0_{t_h,l}\xi_{l,j,h}\right) = \\ &= \mathbb{E}\left[\sum_{k,h=0}^{N} S^0_{t_k,i}S^0_{t_h,l}\xi_{i,j,k}\xi_{l,j,h}\right] - \mathbb{E}\left[\sum_{k=0}^{N} S^0_{t_k,i}\xi_{i,j,k}\right]\mathbb{E}\left[\sum_{h=0}^{N} S^0_{t_h,l}\xi_{l,j,h}\right] \\ &= \sum_{h,k=0}^{N}\xi_{i,j,k}\xi_{l,j,h}\mathrm{Cov}(S^0_{t_k,i},S^0_{t_h,l}) = \sum_{h,k=0}^{N}\xi_{i,j,k}\xi_{l,j,h}\mathrm{Cov}(S^0_{t_k\wedge t_h,i},S^0_{t_k\wedge t_h,l}) \end{aligned}$$

which is zero if the components of $\boldsymbol{S}^0_\cdot$ are uncorrelated. □

Remark 7.7. We observe that for risk-neutral agents, i.e., $\gamma = 0$, the assumptions of uncorrelated assets is no more necessary to prove Lemma 7.6. Indeed, the mean-variance functional is restricted only to the expected cost and for linearity $MV_0(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot}) = \sum_{i=1}^{M} MV_0(\boldsymbol{\xi}_{i,j,\cdot}|\Xi_{i,-j,\cdot};G_i)$, where $MV_0(\boldsymbol{\xi}_{i,j,\cdot}|\Xi_{i,-j,\cdot};G_i) = \mathbb{E}[C_{\mathbb{T}}(\boldsymbol{\xi}_{i,j,\cdot}|\Xi_{i,-j,\cdot};G_i)]$ is the expected costs of Definition 6.5 where the decay kernel is multiplied by λ_i, and we have the same conclusion of Lemma 7.6 regardless the covariance matrix of $\boldsymbol{S}^0_t$.

We first introduce some notation and then we state the main results. We say that assets are orthogonal if the corresponding cross-impact matrix is diagonal. Let us consider the spectral decomposition of Q, i.e., $QV = VD$, where V and D are the orthogonal and diagonal matrices containing the eigenvectors and eigenvalues, respectively. Since we assume that Q is a non singular symmetric matrix, then D is diagonal with all elements different from zero. We define the prices of the virtual assets as $\boldsymbol{P}_t := V^T\boldsymbol{S}^\Xi_t$ and we observe that

$$\begin{aligned} \boldsymbol{P}_t &= \boldsymbol{P}^0_t - \sum_{t_k<t} G(t-t_k)\cdot D\cdot V^T\cdot\left(\sum_{j=1}^{J}\boldsymbol{\xi}_{\cdot,j,k}\right) \\ &= \boldsymbol{P}^0_t - \sum_{t_k<t} G(t-t_k)\cdot D\cdot\left(\sum_{j=1}^{J}\boldsymbol{\xi}^P_{\cdot,j,k}\right), \end{aligned} \tag{7.4}$$

where $\boldsymbol{P}^0_t := V^T\boldsymbol{S}^0_t$ and $\boldsymbol{\xi}^P_{\cdot,j,k} := V^T\boldsymbol{\xi}_{\cdot,j,k}$. This last quantity is the strategy of trader j at time step k in the virtual assets, which is admissible for inventory $\boldsymbol{X}^P_j = V^T\boldsymbol{X}_j$, i.e, $\sum_{k=0}^{N}\boldsymbol{\xi}^P_{\cdot,j,k} = \sum_{k=0}^{N}V^T\boldsymbol{\xi}_{\cdot,j,k} = V^T\boldsymbol{X}_j$. The virtual assets are mutually orthogonal

by construction and their corresponding (virtual) decay kernels $G_i(t)$ are obtained as the product of the original decay kernel $G(t)$ and the corresponding eigenvalues λ_i of the cross impact matrix, i.e., the decay kernel associated with the i-th virtual asset is $G_i(t) := G(t) \cdot \lambda_i$. Indeed, from Eq. (7.4) the decay kernel $G(t)$ is multiplied by the eigenvalues of the cross impact matrix for each trading time t_k,

$$G(t-t_k) \cdot D = \begin{bmatrix} G(t-t_k)\lambda_1 & & & \\ & G(t-t_k)\lambda_2 & & \\ & & \ddots & \\ & & & G(t-t_k)\lambda_M \end{bmatrix}.$$

Then, as observed in Remark 7.5, the multi-asset market impact game where each asset is orthogonal to others is equivalent to M one-asset market impact games, i.e., Luo and Schied (2020) models. The (virtual) decay kernels $G_i(t)$ satisfy the assumptions of strictly positive definite kernels as far as $\lambda_i > 0 \ \forall i = 1, 2, \ldots, M$, i.e., Q is positive definite (see also Alfonsi et al. (2016)). If $\mathrm{Cov}(\boldsymbol{S}_t^0) = \Sigma$, then $\mathrm{Cov}(\boldsymbol{P}_t^0) = V^T \Sigma V$. So, if Q and Σ are simultaneously diagonalizable then $\mathrm{Cov}(\boldsymbol{P}_t^0)$ is diagonal, i.e., the components of $\boldsymbol{P}_\cdot^0$ are uncorrelated and by Lemma 7.6 we obtain the associated Nash equilibria $\Xi^{*,P}$, whose components are defined as

$$\boldsymbol{\xi}_{i,j,\cdot}^{*,P} = \overline{X}_{i,\cdot}^P \boldsymbol{v}_i + (X_{i,j}^P - \overline{X}_{i,\cdot}^P)\boldsymbol{w}_i, \quad j = 1, 2, \ldots, J, \quad i = 1, 2, \ldots, M, \tag{7.5}$$

where $\overline{X}_{i,\cdot}^P = \frac{1}{J}\sum_{j=1}^J X_{i,j}^P$ is the average inventory on the i-th virtual asset among the traders and $\boldsymbol{v}_i$ and $\boldsymbol{w}_i$ are the previously defined fundamental solutions of Luo and Schied (2020) for the i-th virtual asset $P_{\cdot,i}$. For them, the decay kernel is given by $G_i(t) = G(t) \cdot \lambda_i$ and the corresponding $\varphi_i(t)$ is given by $\mathrm{Var}(P_{t,i}^0)$. Since, Q and Σ are both symmetric, so diagonalizable, Q and Σ are simultaneously diagonalizable if and only if Q and Σ commute. Therefore, we consider the following assumption.

Assumptions 7.1. The cross-impact matrix, Q, and the covariance matrix of the unaffected price process $\boldsymbol{S}_t^0$, Σ, commute, i.e., $Q\Sigma = \Sigma Q$.

This assumption is frequently made in the literature and approximately valid in real data, e.g., Mastromatteo et al. (2017) makes this assumption on the correlation matrix. The empirical observation that the matrix Q has a large eigenvalue with a corresponding eigenvector with almost constant components (as the market factor) and a block structure with blocks corresponding to economic sectors (as in the correlation matrix) indicates that the eigenvectors of Q and Σ are the same, i.e. that Q and Σ (approximately) commute. Notice also that Gârleanu and Pedersen (2013) propose a model of optimal portfolio execution where the quadratic transaction cost is characterized by a matrix which is

proportional to Σ.

We enunciate the following theorem of existence and uniqueness of Nash equilibrium which extends Theorem 2.4 of Luo and Schied (2020).

Theorem 7.8 (Nash Equilibrium for Multi-Asset and Multi-Agent Market Impact Games)**.** *For any strictly positive definite decay kernel G, time grid $\mathbb{T}$, parameter $\theta, \gamma \geq 0$, initial inventory $X \in \mathbb{R}^{M\times J}$ and symmetric positive definite cross impact matrix Q such that Assumption 7.1 holds, there exists a unique Nash Equilibrium $\Xi^* \in \mathscr{X}_{\text{det}}(X, \mathbb{T})$ for the mean-variance optimization problem and it is given by*

$$\Xi^*_{\cdot,j,\cdot} = V\Xi^{*,P}_{\cdot,j,\cdot}, \quad j = 1, 2, \ldots, J \tag{7.6}$$

where V is the matrix of eigenvectors of Q and $\Xi^{*,P} \in \mathscr{X}_{\text{det}}(X^P, \mathbb{T})$ is the Nash Equilibrium (7.5) *of the corresponding orthogonalized virtual asset market impact game where $X^P = V^T X$. Moreover, if $\boldsymbol{S}^0$ follows a Bachelier model then* (7.6) *is also a Nash equilibrium for CARA expected utility maximization.*

Proof of Theorem 7.8. Let $Q = VDV^T$ be the spectral decomposition of Q, where, since Q is symmetric, V is orthogonal and D is the diagonal matrix which contains the eigenvalues of Q. By Assumptions 7.1 $\text{Cov}(\boldsymbol{P}^0_t) = V^T\Sigma V$ is diagonal, so by Lemma 7.6 there exists the Nash Equilibrium $\Xi^{*,P} \in \mathscr{X}_{\text{det}}(X^P, \mathbb{T})$, for each inventory X^P associated to the orthogonalized virtual assets $\boldsymbol{P}_t = V^T\boldsymbol{S}_t$. Moreover, if $\boldsymbol{S}^0_t$ follows a Bachelier model then also $\boldsymbol{P}^0_t$ follows a Bachelier model and $\Xi^{*,P}$ is also a Nash equilibrium for the CARA expected utility maximization for Lemma 7.6. Therefore, to proof that Ξ^*, where $\Xi^*_{\cdot,j,\cdot} = V\Xi^{*,P}_{\cdot,j,\cdot}$, is the Nash Equilibrium is sufficient to show that the liquidation cost $C_{\mathbb{T}}(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot})$, when the cross impact matrix is Q, is equivalent to $C_{\mathbb{T}}(\Xi^P_{\cdot,j,\cdot}|\Xi^P_{\cdot,-j,\cdot})$, when the cross impact is D, where the equivalence map is provided by V^T. Writing explicitly for each trading time step k the liquidation cost formula we have, since V is orthogonal,

$$\begin{aligned} C_{\mathbb{T}}(\Xi_{\cdot,j,\cdot}|\Xi_{\cdot,-j,\cdot})] = \sum_{k=0}^{N} \Bigg(& \frac{G(0)}{2}\langle Q\boldsymbol{\xi}_{\cdot,j,k}, \boldsymbol{\xi}_{\cdot,j,k}\rangle - \langle \boldsymbol{S}^{\Xi}_{t_k}, \boldsymbol{\xi}_{\cdot,j,k}\rangle + \\ & + \frac{G(0)}{2}\sum_{l\neq j}\langle Q\boldsymbol{\xi}_{\cdot,l,k}, \boldsymbol{\xi}_{\cdot,j,k}\rangle + \theta\ \langle \boldsymbol{\xi}_{\cdot,j,k}, \boldsymbol{\xi}_{\cdot,j,k}\rangle_{\cdot}\Bigg) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=0}^{N} \left(\frac{G(0)}{2} \langle DV^T \boldsymbol{\xi}_{\cdot,j,k}, V^T \boldsymbol{\xi}_{\cdot,j,k} \rangle - \langle V^T \boldsymbol{S}^{\Xi}_{t_k}, V^T \boldsymbol{\xi}_{\cdot,j,k} \rangle + \right. \\
&+ \left. \frac{G(0)}{2} \sum_{l \neq j} \langle DV^T \boldsymbol{\xi}_{\cdot,l,k}, V^T \boldsymbol{\xi}_{\cdot,j,k} \rangle + \theta \, \langle V^T \boldsymbol{\xi}_{\cdot,j,k}, V^T \boldsymbol{\xi}_{\cdot,j,k} \rangle_{\cdot} \right) \\
&= \sum_{k=0}^{N} \left(\frac{G(0)}{2} \langle D\Xi^P_{\cdot,j,k}, \Xi^P_{\cdot,j,k} \rangle - \langle \boldsymbol{P}_{t_k}, \Xi^P_{\cdot,j,k} \rangle + \right. \\
&+ \left. \frac{G(0)}{2} \sum_{l \neq j} \langle D\Xi^P_{\cdot,l,k}, \Xi^P_{\cdot,j,k} \rangle + \theta \, \langle \Xi^P_{\cdot,j,k}, \Xi^P_{\cdot,j,k} \rangle_{\cdot} \right) \\
&= C_{\mathbb{T}}(\Xi^P_{\cdot,j,\cdot} | \Xi^P_{\cdot,-j,\cdot}).
\end{aligned}
$$

Finally, in order to obtain that Ξ^* is admissible for X, it is sufficient to set $X^P = V^T X$. □

However, we observe that for risk-neutral agents, i.e., $\gamma = 0$, Assumption 7.1 is unnecessary. We remark this result in the following Corollary.

Corollary 7.9. *If the agents are risk-neutral, i.e., $\gamma = 0$, then for any strictly positive definite decay kernel G, time grid $\mathbb{T}$, parameter $\theta \geq 0$, initial inventories $X \in \mathbb{R}^{M \times J}$ and symmetric positive definite cross impact matrix Q, there exists a unique Nash Equilibrium $\Xi^* \in \mathscr{X}_{\det}(X, \mathbb{T})$ for the mean-variance optimization problem and it is given by*

$$
\Xi^*_{\cdot,j,\cdot} = V \Xi^{*,P}_{\cdot,j,\cdot}, \quad j = 1, 2, \ldots, J \tag{7.7}
$$

where V is the matrix of eigenvectors of Q and $\Xi^{*,P} \in \mathscr{X}_{\det}(X^P, \mathbb{T})$ is the Nash Equilibrium associated to the corresponding orthogonalized virtual asset market impact game where $X^P = V^T X$. Moreover, if $\boldsymbol{S}^0_t$ follows a Bachelier model then (7.7) is also a Nash equilibrium over the set $\mathscr{X}(X, \mathbb{T})$.

Proof of Corollary 7.9. As observed in Remark 7.7 the mean-variance functional is splitted as the sum of mean-variance functionals of each asset i, since when $\gamma = 0$ the functional is restricted to the expected cost. Then, the existence of the Nash equilibrium for the virtual orthogonalized assets follows by Lemma 7.6 without requiring the assumptions of uncorrelated assets and the proof follows directly by the same reasoning of the proof of Theorem 7.8. Moreover by definition, when $\gamma = 0$ the CARA utility function is equal to the mean-variance functional, so that Ξ^* is a Nash equilibrium over the set $\mathscr{X}(X, \mathbb{T})$. □

CHAPTER 8

Trading Strategies in Market Impact Games

Before studying market stability we investigate how the cross-impact effect and the presence of many competitors may affect trading strategies, in terms of Nash equilibria. To understand the rich phenomenology that can be observed in a market impact game, we introduce three types of traders:

- the *Directional (Fundamentalist)* wants to trade one or more assets in the same direction (buy or sell). Notice that a Directional can have zero initial inventory for some assets;
- the *Arbitrageur* has a zero inventory to trade in each asset and tries to profit from the market impact payed by the the other agents;
- the *Market Neutral* has a non zero volume to trade in each asset, but in order to avoid to be exposed to market index fluctuations, the sum of the volume traded in all assets is zero[1].

We remark that an Arbitrageur is a particular case of a Market Neutral agent in the limit case when the volume to trade in each asset is zero. Clearly in a single-asset market we have only two types of the previous agents, since a Market Neutral strategy requires at least two assets.

8.1 Cross-Impact effect and liquidity strategies

To better understand how cross-impact affects optimal liquidation strategies, we consider the case of two risk-neutrals agents which can (but not necessarily mist) trade M assets. We show below that the presence of multiple assets and of cross-impact can affect the trading strategy of an agent interested in liquidating only one asset. In particular, we find, counterintuitively, that it might be convenient for such an agent to trade (with zero inventory) the other asset(s) in order to reduce transaction costs.

[1]Real Market Neutral agents follow signals which are orthogonal to the market factor, thus they typically are short on approximately half of the assets and long on the other half. The sum of trading volume is not exactly equal to zero but each trading volume depends on the β of the considered asset with respect to the market factor. In our stylized market setting, we assume that all assets are equivalent with respect to the market factor.

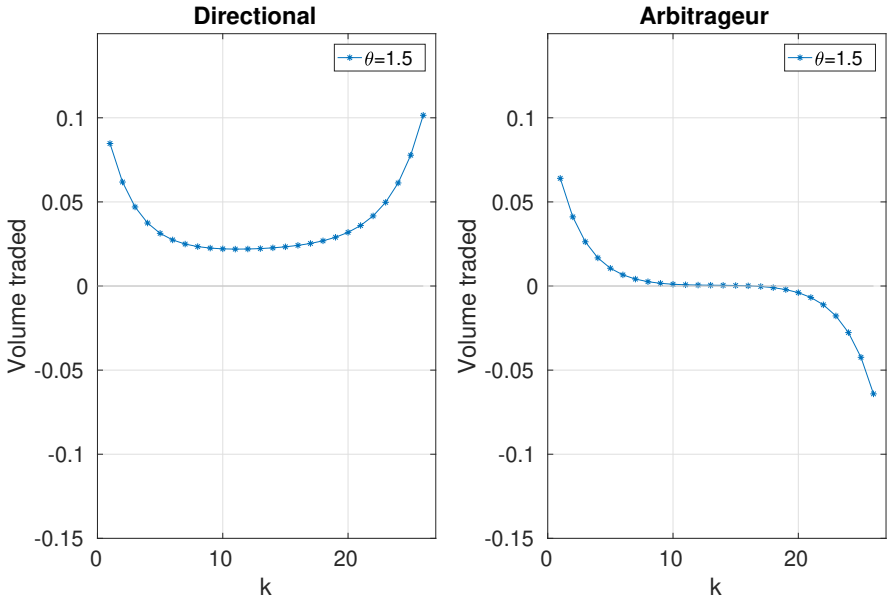


Fig. 8.1 Nash equilibrium $\boldsymbol{\xi}_1^*$ of the Directional and $\boldsymbol{\xi}_2^*$ of the Arbitrageur trading only one asset. The trading time grid is equidistant with 26 points and $\theta = 1.5$. The expected costs are equal to $\mathbb{E}[C_{\mathbb{T}}(\boldsymbol{\xi}_1^*|\boldsymbol{\xi}_2^*)] = 0.4882$, $\mathbb{E}[C_{\mathbb{T}}(\boldsymbol{\xi}_2^*|\boldsymbol{\xi}_1^*)] = -0.0370$.

We focus on the two-asset case, $M = 2$, and we analyze the Nash equilibrium when the kernel function has an exponential decay[2], $G(t) = e^{-t}$. The first trader is a Directional who wants to liquidate the position in the first asset, i.e., $X_{1,1} = 1$, while the second agent is an Arbitrageur, i.e., $X_{1,2} = 0$. We set an equidistant trading time grid with 26 points and $\theta = 1.5$. The second asset is available for trading, but let us consider as a benchmark case when both agents trade only the first asset. This is a standard Schied and Zhang (2018) game. Figure 8.1 exhibits the Nash Equilibrium for the two players. We observe that the optimal solution for the Directional is very close to the classical U-shape derived under the Transient Impact Model (TIM)[3], i.e., our model when only one agent is present. However, the solution is asymmetric and it is more convenient for the Directional to trade more in the last period of trading. This can be motivated by observing that at equilibrium the Arbitrageur places buy order at the end of the trading day, and thus she pushes up the price. Then, the Directional exploits this impact to liquidate more orders at the end of the trading session. We remark that the Arbitrageur earns at equilibrium, since her expected cost is negative (see the caption).

Now we examine the previous situation when the two traders solve the optimal execution problem taking into account the possibility of trading the other asset. We define the

[2]All our numerical experiments are performed with exponential kernel as in (Obizhaeva and Wang (2013)). Schied and Zhang shows that the form of the kernel does not play a key role for stability, given that the conditions of Proposition 6.4 are satisfied.

[3]Given the initial inventory X, the optimal strategy in the standard TIM is $\boldsymbol{\xi} = \frac{X}{\boldsymbol{e}^T\Gamma_\theta^{-1}\boldsymbol{e}}\Gamma_\theta^{-1}\boldsymbol{e}$, see for further details Schied and Zhang (2018).

Table 8.1 Payoff matrix of expected costs when the Directional and Arbitrageur inventories are equal to $(1\ 0)^T$ and $(0\ 0)^T$, respectively. We have highlighted in red the Nash Equilibrium associated with this payoff matrix. The payoff in the i-th row and j-th column correspond to the game when the Directional and Arbitrageur decide to trade i and j assets, respectively, i.e., the element in the first row and second column is the payoff when the Directional trades only the first asset while the Arbitrageur trades both assets.

		Arbitrageur	
		1 Asset	2 Asset
Directional	1 Asset	$(0.4882, -0.0370)$	$(0.4935, -0.0412)$
	2 Asset	$(0.4836, -0.0334)$	$(0.4885, -0.0377)$

cross impact matrix $Q = \begin{bmatrix} 1 & q \\ q & 1 \end{bmatrix}$, where $q = 0.6$. In Figure 8.2 we report the optimal solution where the inventory of the agents are set to be $\boldsymbol{X}_1 = \begin{pmatrix} 1 & 0 \end{pmatrix}^T$ and $\boldsymbol{X}_2 = \begin{pmatrix} 0 & 0 \end{pmatrix}^T$. The Directional wants to liquidate only one asset, but, as clear from the Nash equilibrium, the cross-impact influences the optimal strategies in such a way that it is optimal for him/her to trade also the other asset. In terms of cost, for the Directional trading the two assets is worse off than in the benchmark case (see the values of $\mathbb{E}[C_{\mathbb{T}}(\Xi^*_{\cdot,1,\cdot}|\Xi^*_{\cdot,2,\cdot})]$ in captions). However, if the Directional trades only asset 1 and Arbitrageur trades both assets, the former has a cost of 0.4935 which is greater than the expected costs associated with Figure 8.2. Thus, the Directional *must* trade the second asset if the Arbitrageur does (or can do it).

For completeness in Table 8.1 we compare the expected costs of both Directional and Arbitrageur when the two agents may decide to trade i) both assets, i.e., they consider market impact game and cross-impact effect, or ii) one asset, i.e., they only consider the market impact game. It is clear that both agents prefer to trade both assets. Actually, the state where both agents trade two assets is the Nash equilibrium of the game where each agent can choose how many assets to trade.

The solution presented above is generic, but an important role is played by the transaction cost modeled by the temporary impact. When the temporary impact parameter θ increases, the benefit of the cross-impact vanishes, and the optimal strategy of the Directional tends to the solution provided by the simple TIM with one asset and no other agent. We find that the difference between these expected costs is negative, i.e. it is always optimal to trade also the second asset, but converges to zero for large θ, see Figure 8.3 panel (a). Furthermore, it is worth noting that, if $S = \sum_k |\xi_{k,2}|$ denotes the total absolute volume traded by the Directional on the second asset, then $\lim_{\theta\to 0} S = 0$ and $\lim_{\theta\to\infty} S = 0$ as exhibited from Figure 8.3 panel (b). This means, that when the cost of trades increases, it is not anymore convenient for both traders to try to exploit the cross impact effect.

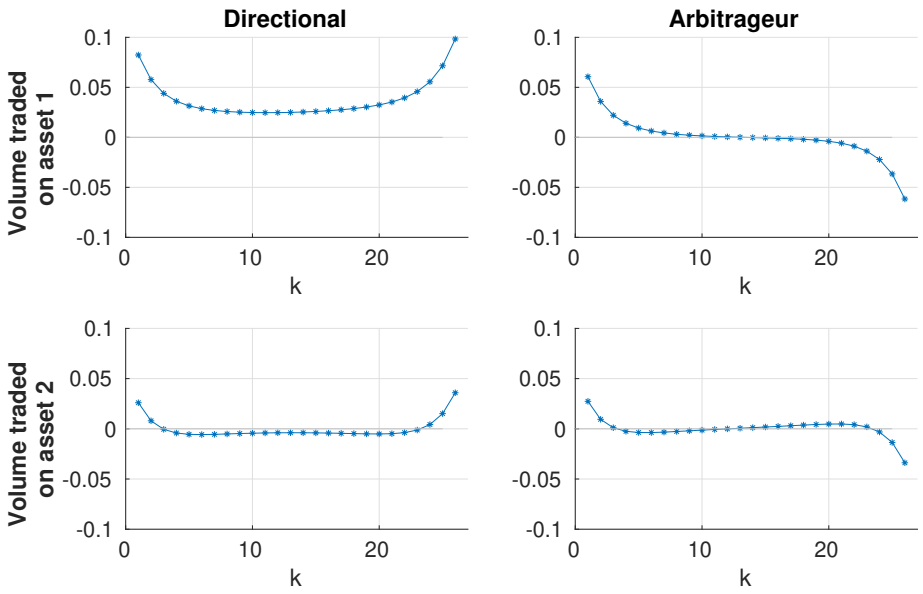


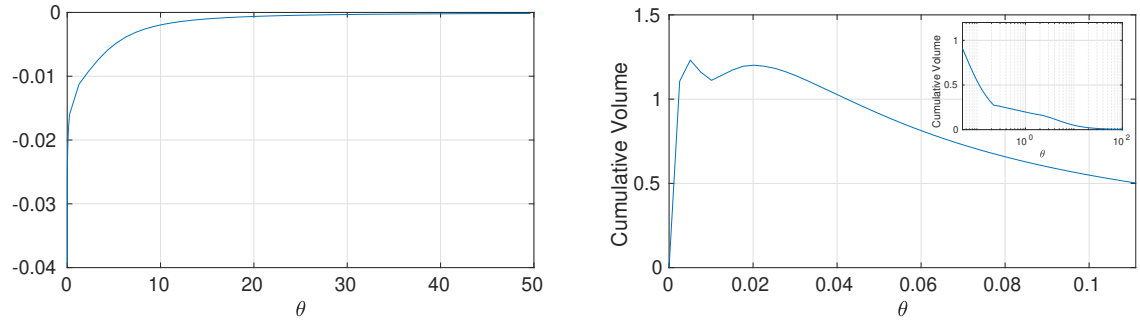
Fig. 8.2 Optimal strategies for a Directional ($\Xi^*_{\cdot,1,\cdot}$) and an Arbitrageur ($\Xi^*_{\cdot,2,\cdot}$), where their inventories are equal to $(1\ 0)^T$ and $(0\ 0)^T$, respectively. $Q = \begin{bmatrix} 1 & 0.6 \\ 0.6 & 1 \end{bmatrix}$, and the trading time grid is an equidistant time grid with 26 points. The expected costs are equal $\mathbb{E}[C_{\mathbb{T}}(\Xi^*_{\cdot,1,\cdot}|\Xi^*_{\cdot,2,\cdot})] = 0.4885$, $\mathbb{E}[C_{\mathbb{T}}(\Xi^*_{\cdot,2,\cdot}|\Xi^*_{\cdot,1,\cdot})] = -0.0377$ when $\theta = 1.5$.

8.2 Do arbitrageurs act as market makers at equilibrium?

We now consider the cases when the agents are of different type. In particular, we focus on the role of an Arbitrageur as an intermediary between two Fundamental traders of opposite sign. When a Fundamental seller and a Fundamental buyer trade the same asset(s), are the Arbitrageurs able to profit, acting as a sort of market maker by buying from the former and selling to the latter?

To answer this question, we compute the Nash equilibrium of a market impact game with $M = 2$ assets and $J = 3$ agents, namely a Directional seller with inventory $(1\ 0)^T$, a Directional buyer with inventory $(-1\ 0)^T$, and an Arbitrageur. We assume that agents are risk-neutrals, $\gamma = 0$, and $Q = \begin{bmatrix} 1 & 0.6 \\ 0.6 & 1 \end{bmatrix}$. As panels (a) of Figure 8.4 show, the Arbitrageur does not longer trade and the expected costs are 0.1056 and 0 for the two Directionals and the Arbitrageur, respectively. This indicates that the two Directionals are able to reduce significantly their costs with respect to the previous case, increasing their protection against predatory trading strategies and that the Arbitrageur is unable to act as a market maker. The previous cases are particular examples of the following more general result.

Proposition 8.1. *Under the assumptions of Theorem 7.8, the following are equivalent:*



(a) **(b)**

Fig. 8.3 Figure (a). The y axis shows the difference between the expected cost of the Directional when he/she consider the cross-impact effect and the Arbitrageur and the expected cost when he/she places order following the classical one asset TIM model and the x axis the cost parameter θ. Figure (b). Cumulative traded volume of the second asset by the Directional when playing against an Arbitrageur as a function of θ. The inset shows the same curve in semi-log scale. The setting is the same of Figure 8.2.

a) The aggregate net order flow is zero for each asset, i.e.,

$$\overline{X}_{i,\cdot} = \frac{1}{J}\sum_{j=1}^{J} X_{i,j} = 0 \quad \forall i = 1, 2, \ldots, M;$$

b) The optimal solution for an Arbitrageur is equal to zero for all assets.

Proof of Proposition 8.1. Let the j-th trader be an Arbitrageur, i.e., $\boldsymbol{X}_{\cdot,j} = \mathbf{0} \in \mathbb{R}^M$. Moreover, his/her inventory for the virtual assets is zero, $X^P_{i,j} = \sum_{m=1}^{M} V^T_{i,m} X_{m,j} = 0$ for each $i = 1, 2, \ldots, M$. Then, since for Theorem 7.8 Eq. (7.5) provides the optimal schedule on each virtual assets i, the optimal schedule of the Arbitrageur for the i-th virtual asset is characterized by the corresponding $\overline{X}^P_{i,\cdot}$.

a) $\Rightarrow$ b). If $\overline{X}_{i,\cdot} = 0, \ \forall i$ then

$$\overline{X}^P_{i,\cdot} = \frac{1}{J}\sum_{j=1}^{J} X^P_{i,j} = \frac{1}{J}\sum_{j=1}^{J}\sum_{m=1}^{M} V^T_{i,m} X_{m,j} = \sum_{m=1}^{M} V^T_{i,m}\overline{X}_{m,\cdot} = 0 \quad \forall i.$$

So, the solution of the Arbitrageurs for each virtual assets is zero and hence also for the original assets by Theorem 7.8.

b) $\Rightarrow$ a). If the optimal solution for an Arbitrageur is zero for all assets, then by Theorem 7.8 and since V is orthogonal, the optimal solution for the Arbitrageur is zero also for the virtual assets, so that $\overline{X}^P_{i,\cdot} = 0 \ \forall i$ and then $\overline{X}_{i,\cdot} = 0 \ \forall i$. □

In other words, when the aggregate net order flow is zero for each asset then there are no arbitrageurs in the market, i.e., the Nash equilibrium for Arbitrageurs is *zero*, so that the optimal schedule corresponds to place no orders in the market.

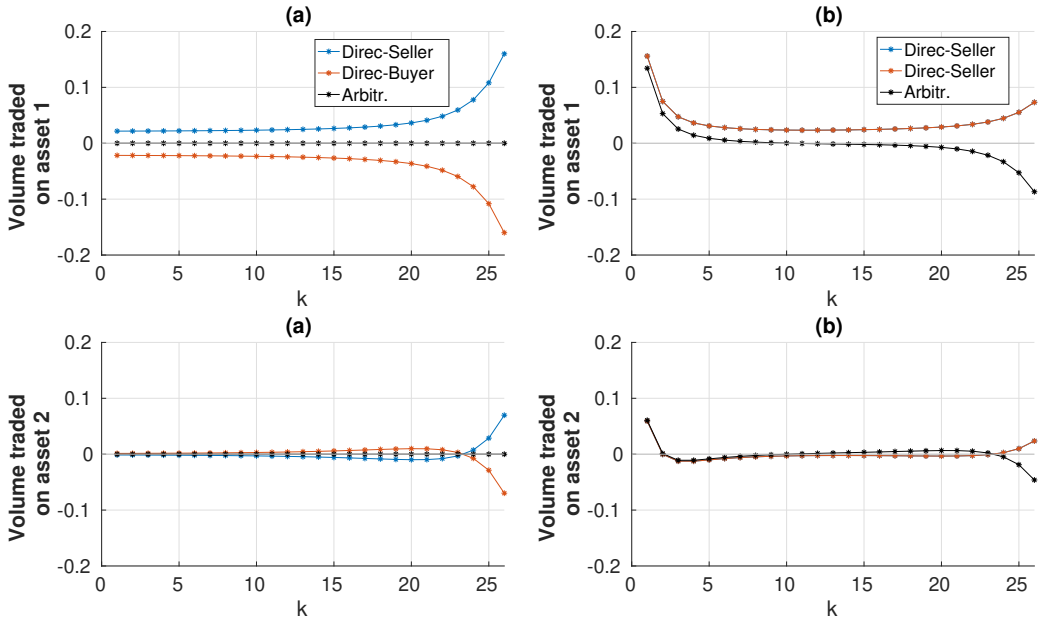


Fig. 8.4 Optimal schedule for market impact game with $M = 2$ assets and $J = 3$ risk-neutral agents. Panels (a) exhibit the optimal schedule for a Directional seller, buyer (with inventory $(1\ 0)^T$ and $(-1\ 0)^T$, respectively), and an Arbitrageur. Panels (b) exhibit the optimal schedule for two identical Directional sellers (with inventories $(1\ 0)^T$, respectively), and an Arbitrageur. Blue and red lines are the Nash equilibrium for the Directional traders. The black line refers to the equilibrium of the Arbitrageur. The trading time is equidistant with 26 points, where the cross impact is set to $q = 0.6$, $\gamma = 0$ and $\theta = 1.5$.

As a comparison, we consider two identical Directional sellers (with inventories $(1\ 0)^T$) and the other parameters are the same as above. Figure 8.4, panels (b), displays the equilibrium solution. The solution of the Directionals are identical. While the trading pattern of the Arbitrageur is qualitatively similar to the one of the two agent case (see Fig. 8.2), the Directionals trade significantly less toward the end of the day. This is likely due to the fact that it might be costly to trade for one Directional given the presence of the other. The expected costs of the two Directionals is equal to 0.8911 (which is approximately two times of the two players game) and -0.0996 for the Arbitrageur.

CHAPTER 9

Instability in Market Impact Games

We now turn to our attention to the study of market stability. Since the seminal work of Schied and Zhang (2018) we known that, when two risk-neutral agents trade one asset, stability is fully determined by the behavior at the origin of the decay kernel, see Theorem 2.7 of Schied and Zhang (2018). Here we extend their results for the multi-asset case and we derive a general result which involves the spectrum of the cross-impact matrix. However, the proof of Schied and Zhang (2018) cannot[1] be extended to the multi-agent case with J risk-averse agents, even though in the one asset case, as highlighted by Luo and Schied (2020). Therefore, we study market stability by using numerical analyses for the general setting of multi-agent and multi-asset case from which we deduce a new conjecture which is in line with the analyses carried out by Luo and Schied (2020). We conclude by presenting some advice to policy regulators which want to prevent market instability.

To clarify better our results, we introduce two definitions of market stability in a market with M assets and J traders:

Definition 9.1 (Strong Stability)**.** The market is *strongly (uniformly) stable* if $\forall\ \theta \geq 0$ the Nash equilibrium $\boldsymbol{\xi}^*_{i,j,\cdot} \in \mathscr{X}(X_{i,j}, \mathbb{T})$ does not exhibit spurious oscillations $\forall\ X_{i,j} \in \mathbb{R}$ initial inventory, for all assets $i = 1, 2, \ldots, M$ and agents $j = 1, 2, \ldots, J$.

Definition 9.2 (Weak Stability)**.** The market is *weakly stable* if there exists an interval $I \subset \mathbb{R}_+$ such that $\forall\ \theta \in I$ the Nash equilibrium $\boldsymbol{\xi}^*_{i,j,\cdot} \in \mathscr{X}(X_{i,j}, \mathbb{T})$ does not exhibit spurious oscillations $\forall\ X_{i,j} \in \mathbb{R}$ initial inventory, for all assets $i = 1, 2, \ldots, M$ and agents $j = 1, 2, \ldots, J$

We recall that a spurious oscillations is a sequence of trading times such that the orders are consecutively composed by buy and sell trades, see Chapter 6. Therefore, Schied and Zhang (2018) showed that for $M = 1$ and $J = 2$ the market is not strongly but only weakly stable where I, the stability region, is equal to $[\theta^*, +\infty)$ where $\theta^* = G(0)/4$.

[1]The proofs provided of Schied and Zhang (2018) rely on general results of Toeplitz matrix, which cannot be used in the multi-agent framework, since the involved decay kernel matrices are no longer Toeplitz.

9.1 Market stability and cross impact structure

In this Section we consider $J = 2$ risk-neutral agents which trade $M > 1$ assets. We study whether the increase of the number of assets and the structure of cross impact matrix help avoiding oscillations and market instability at equilibrium according to the previous definitions. To this end, we consider different structures of the cross-impact matrix Q describing the complexity of the market for what concerns commonality in liquidity.

We first show that instabilities are generically observed also in the multi-asset case and that actually more assets generally make the market less stable. For simplicity let us consider $M = 2$ assets and a game between a Directional and an Arbitrageur (similar results hold for different combinations of agents). We choose $G(t) = e^{-t}$, the cross impact matrix equal to $Q = \begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix}$, and we consider $\theta = 0.3$; remember that for the one asset case the market is stable for this value of θ. Figure 9.1 shows that for this value of θ the strategies are oscillating and therefore the market is not strongly stable. More surprisingly, the fact that oscillations are observed for $\theta = 0.3$ indicates that the transition between the two stability regimes depends on also on the number of assets and that more assets require larger values of θ to ensure stability. In the following we prove that this is the case and we determine the threshold value. Figure 9.1 shows also the case $\theta = 0$. Notably, in this case the oscillations in the second asset disappear. This is due to the fact that, since Γ_0^1, (Γ_0^2), the Γ matrix associated with the first (second) virtual asset is equal to $(1+q)\Gamma$, $((1-q)\Gamma)$, the combination of "fundamental" solutions $\boldsymbol{v}$ and $\boldsymbol{w}$ are the same for the two virtual assets. Thus, at equilibrium the two solutions for the second asset are exactly zero.

We have shown in a simple setting that having more than one available asset does not help improving the strong stability of the market and increases the threshold value between stable and unstable markets. Now, we show that when the number of assets tends to infinity the market does not satisfy the weak stability condition. Indeed, in the one asset setting, if we choose a sufficiently large θ the instability vanishes. Therefore, this raises the question of whether the equilibrium instability is still present when the number of assets increases. To this end we introduce the definition of asymptotic stability.

Definition 9.3 (Asymptotically weakly stable)**.** The market is asymptotically weakly stable if it is weakly stable when $M \to \infty$.

Given this definition, we prove the following:

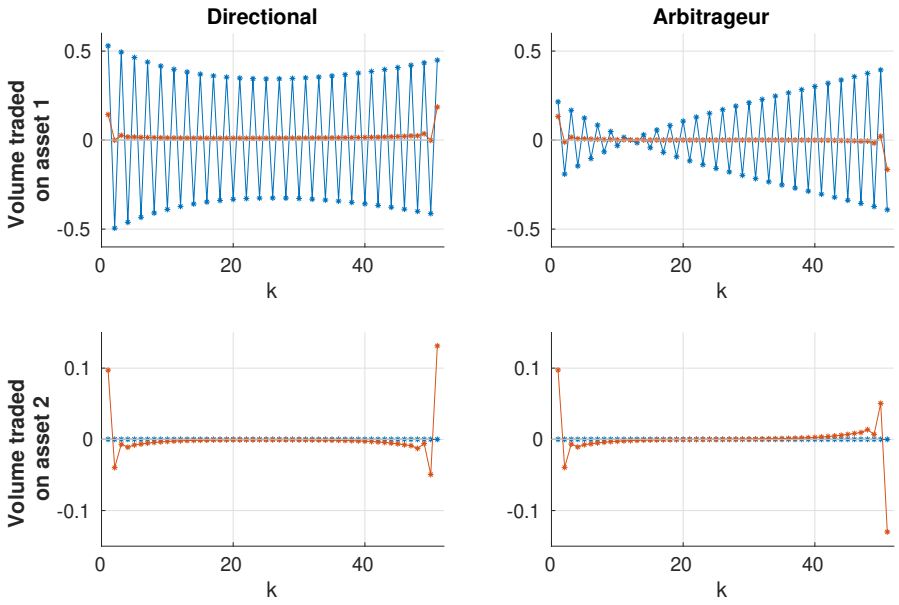


Fig. 9.1 Nash Equilibrium for a Directional and an Arbitrageur, where their inventories are equal to $(1\ 0)^T$ and $(0\ 0)^T$ respectively. The blue lines are the optimal solution when $\theta = 0$ and the red lines when $\theta = 0.3$. The trading time has 51 points and $Q = \begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix}$.

Theorem 9.4 (Instability in Multi-Asset Market Impact Games)**.** *Suppose that G is a continuous, positive definite, strictly positive, log-convex decay kernel and that the time grid is equidistant. Let $(\lambda_i)_{i=1,..,M}$ be the eigenvalues of the cross-impact matrix Q. If $\theta < \theta^*$ the market is unstable, where*

$$\theta^* = \max_{i=1,2,...,M} \frac{G(0) \cdot \lambda_i}{4}. \tag{9.1}$$

Proof of Theorem 9.4. Let $\boldsymbol{X}_1, \boldsymbol{X}_2$ be the inventories of trader first and second trader, respectively. In order to show that market is unstable it is sufficient to exhibit initial inventories which lead to optimal trading strategies with spurious oscillations. WLOG we may assume that inventories are normalized to 1, i.e., $\boldsymbol{X}_1^T \boldsymbol{X}_1 = \boldsymbol{X}_2^T \boldsymbol{X}_2 = 1$. Therefore, let us consider $\boldsymbol{X}_1 = -\boldsymbol{X}_2$, so that $\boldsymbol{X}_1^P = V^T \boldsymbol{X}_1 = -V^T \boldsymbol{X}_2 = -\boldsymbol{X}_2^P$ and the NE for the i-th virtual assets is fully characterized by the fundamental solutions $\boldsymbol{w}_i$. So, for each virtual asset the instability is led by the correspondent virtual kernel, i.e., the kernel relative to the i-th virtual asset which is given by $G \cdot \lambda_i$, where λ_i is the related i-th eigenvalues. Then, for the Schied and Zhang instability result we know that if we want non oscillatory solutions, θ has to be greater than $G(0) \cdot \lambda_i / 4$ for all i. However, if $\boldsymbol{\nu}_i$ denotes the i-th eigenvector of Q, which may be assumed normalized $\boldsymbol{\nu}_i^T \boldsymbol{\nu}_i = 1$, then when $\boldsymbol{X}_1 = \boldsymbol{\nu}_i$ the virtual inventory $\boldsymbol{X}_1^P$ has 1 in the i-th component and zero otherwise.

Then, $\Xi^{*,P}$ is a matrix where the i-th row is equal to $\boldsymbol{w}_i^T$ and zero otherwise. Therefore,

$$\Xi^* = V \cdot \Xi^{*,P} = \left[\boldsymbol{\nu}_1 | \cdots \boldsymbol{\nu}_{i-1} | \boldsymbol{\nu}_i | \boldsymbol{\nu}_{i+1} | \cdots | \boldsymbol{\nu}_M\right] \cdot \Xi^{*,P} = \begin{bmatrix} \nu_{1,i}\boldsymbol{w}_i^T \\ \vdots \\ \nu_{M,i}\boldsymbol{w}_i^T \end{bmatrix} = \boldsymbol{\nu}_i \otimes \boldsymbol{w}_i,$$

i.e., the NE for the j-asset is given by $v_{j,i}\boldsymbol{w}_i$, so also the stability for the original asset $\boldsymbol{S}_t$ is characterized by $\boldsymbol{w}_i$. Then, if $\theta < \theta^* = \max_{i=1,2,\ldots,M} \frac{G(0)\cdot\lambda_i}{4}$ and $i_{\max}$ denotes the position of the maximum eigenvalue, the NE for inventories $\boldsymbol{X}_1 = -\boldsymbol{X}_2 = \boldsymbol{\nu}_{i_{\max}}$ exhibits spurious oscillations. □

Moreover, if the largest eigenvalue of the cross-impact matrix diverges for $M \to \infty$, i.e., $\lim_{M\to+\infty} \lambda_{max} = +\infty$, then the market is not asymptotically weakly stable. The theorem tells that the instability of the market is related to the spectral decomposition of the cross-impact matrix, i.e., to the liquidity factors.

We analyze some realistic cross-impact matrices and their implications for the stability of the Nash equilibrium. Schneider and Lillo (2019) have derived constraints on the structure of the cross-impact for the absence of dynamic arbitrage. They showed that the symmetry of the cross-impact matrix is one of these conditions. Mastromatteo et al. (2017) estimated the cross-impact matrix on 150 US stocks showing that it is roughly symmetric and has a block structure with blocks related to economic sectors. Specifically, we consider the one-factor and block matrices.

9.1.1 One factor matrix

We say that Q is a one factor matrix if $Q = (1-q)I + q \cdot \boldsymbol{e}\boldsymbol{e}^T$, where $\boldsymbol{e} = (1, \ldots, 1)^T \in \mathbb{R}^M$ and $q \in (0, 1)$. The bounds on q guarantee the positive definiteness of the cross-impact matrix. Then it holds:

Corollary 9.5. *Under the assumptions of Theorem 9.4, if the cross-impact matrix is a one factor matrix, then the market is not asymptotically weakly stable.*

Proof of Corollary 9.5. The eigenvalues of Q are $\lambda_1 = 1 - q + qM$ and $\lambda_{2:M} = 1 - q$, where $\boldsymbol{v}_1 = \boldsymbol{e}$, the vector with all 1, is the virtual asset associated with λ_1. Then, when $M \to \infty$ the first eigenvalue diverges so for Theorem 9.4 we conclude. □

This implies that when M increases the transactions cost θ must raise in order to prevent market instability, since $\theta^* = G(0)\lambda_{max}/4 \sim G(0)qM/4$, because $\lambda_{max} = 1 + q(M-1)$.

Figure 9.2 exhibits the equilibrium for a Directional and an Arbitrageur, when $\theta = 1.5$, $q = 0.2$ and $M = 2000$. The inventory of the Directional is 1 for the first 1000 assets and zero for the others. The solutions clearly show spurious oscillations of buy and sell

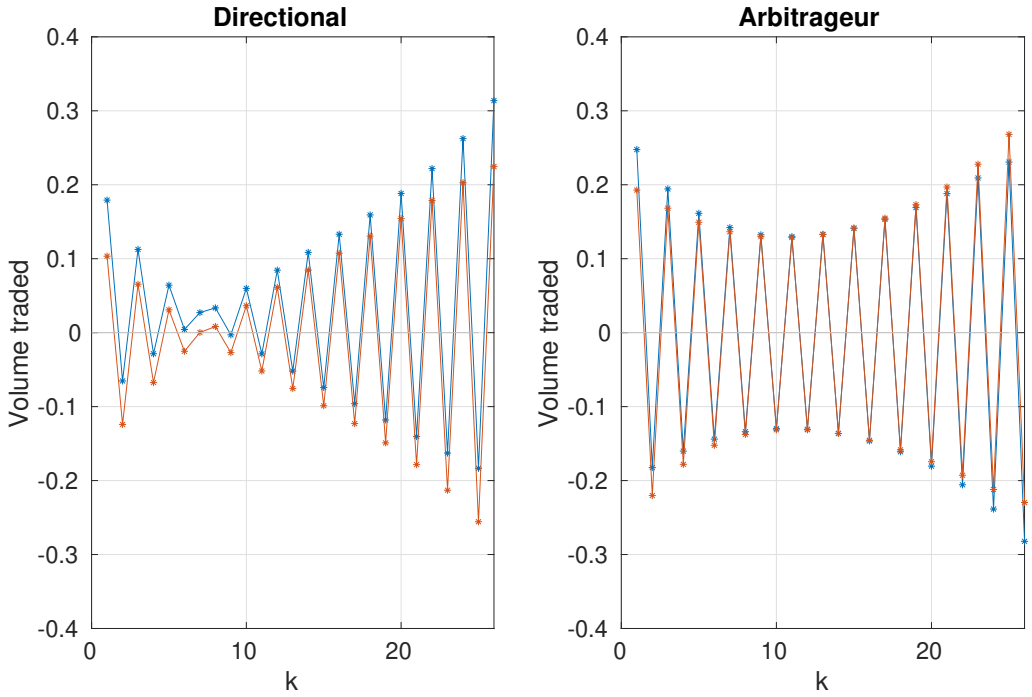


Fig. 9.2 Nash equilibrium when $\theta = 1.5$ between a Directional with inventory $(1, \dots, 1, 0, \dots, 0)^T \in \mathbb{R}^M$ and an Arbitrageur with inventory $(0, \dots, 0)^T \in \mathbb{R}^M$, where $M = 2,000$. The cross impact matrix is a one factor matrix with $q = 0.2$. The blue lines exhibits the volume traded for any of the first $1,000$ assets, while the red ones are those for any of the last $1,000$ assets. The equidistant time grid has 26 points.

orders. Notice that in the one asset case this value of θ gives a stable market. We observe that the eigenvector corresponding to λ_{max} is given by $\boldsymbol{e}$, which represents an equally weighted portfolio. As a consequence, if we consider a Market Neutral agent against an Arbitrageur the solution becomes stable $\forall\ \theta > (1-q)/4$, since both traders have zero inventory on the first virtual asset. Thus, oscillations might disappear when the inventory of the agents in the first virtual asset is zero.

A generalization of the above model considers Q as a rank-one modification matrix, i.e., $Q = D + \boldsymbol{\beta}\boldsymbol{\beta}^T$, where $D = \text{diag}(1-\beta_1^2, \dots, 1-\beta_M^2)$ and $\boldsymbol{\beta} \in \mathbb{R}^M$ is a fixed vector. In this way the cross impact is not the same across all pairs of stocks. We find again that the market is not asymptotically stable because the threshold increases with M. Differently from the previous case this is observed also in the case of a Market Neutral against an Arbitrageur[2].

9.1.2 Block matrix

We now assume that the cross impact matrix has a block structure in such a way that cross impact between two stocks in the same block i is q_i, while when the two stocks

[2]For the sake of simplicity we omit the figure which exhibits the strategies of the two traders and it is available upon request.

are in different blocks the cross impact is q, which we assume to be $0 \leq q < q_i \ \forall i$. As mentioned above, this is consistent with the empirical evidence in Mastromatteo et al. (2017).

Let us denote with M_i the number of stocks in block i, $(i = 1, \dots K)$, and let $Q_i = (1 - q_i)I + q_i \cdot \boldsymbol{e}_i \boldsymbol{e}_i^T \in \mathbb{R}^{M_i} \times \mathbb{R}^{M_i}$ with $q_i \in (0, 1)$ and $\boldsymbol{e}_i = (1, \dots, 1)^T \in \mathbb{R}^{M_i}$, where K is the number of blocks. We define the cross impact matrix as:

$$Q := \begin{bmatrix} Q_1 & q\boldsymbol{e}_1\boldsymbol{e}_2^T & \cdots & q\boldsymbol{e}_1\boldsymbol{e}_K^T \\ q\boldsymbol{e}_2\boldsymbol{e}_1^T & Q_2 & \cdots & q\boldsymbol{e}_2\boldsymbol{e}_K^T \\ \vdots & & \ddots & \vdots \\ q\boldsymbol{e}_K\boldsymbol{e}_1^T & \cdots & q\boldsymbol{e}_K\boldsymbol{e}_{K-1}^T & Q_K \end{bmatrix},$$

If the average number of stocks of a cluster tends to infinity when M goes to infinity, we prove an analogue result as for the one factor matrix case:

Corollary 9.6. *Under the assumptions of Theorem 9.4, if Q is a block matrix, where each block is a one factor matrix, if $\lim_{M\to+\infty} \frac{M}{K} \to +\infty$, then the market is not asymptotically weakly stable.*

Proof of Corollary 9.6. We first note that by Theorem 9.4 it is sufficient to prove that there exists a cluster which is unbounded. Indeed, we observe that

$$Q = \widehat{Q} + q \begin{bmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \\ \vdots \\ \boldsymbol{e}_K \end{bmatrix} \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_K \end{bmatrix}$$

where

$$\widehat{Q} = \begin{bmatrix} Q_1 - q\boldsymbol{e}_1\boldsymbol{e}_1^T & 0 & \cdots & 0 \\ 0 & Q_2 - q\boldsymbol{e}_2\boldsymbol{e}_2^T & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & Q_K - q\boldsymbol{e}_K\boldsymbol{e}_K^T \end{bmatrix}.$$

Then by Theorem 8.1.8 pag.443 of Golub and Van Loan (2013) $\lambda_1(Q) \geq \lambda_1(\widehat{Q})$ where $\lambda_i(Q)$ denotes the i-th largest eigenvalue of Q and respectively of $\widehat{Q}$. However, the eigenvalues of $\widehat{Q}$ are given by the eigenvalues of $Q_i - q\boldsymbol{e_i}\boldsymbol{e_i}^T$ for $i = 1, 2, \dots, K$. For each i, $\lambda_1(Q_i - q\boldsymbol{e_i}\boldsymbol{e_i}^T) = 1 - q_i + M_i(q_i - q)$ and the rests $M_i - 1$ eigenvalues are equal to $1 - q_i$. So, if there exists a cluster such that M_i is unbounded for any value of θ, then $\lambda_1(Q_i - q\boldsymbol{e}_i\boldsymbol{e}_i^T)$ is unbounded and also the respective eigenvalue of Q, so by Theorem 9.4 we conclude that there is no a finite value for θ such that the market is weakly stable.

So, let us first start by fixing the number of cluster to $K < \infty$. Then, when M tends

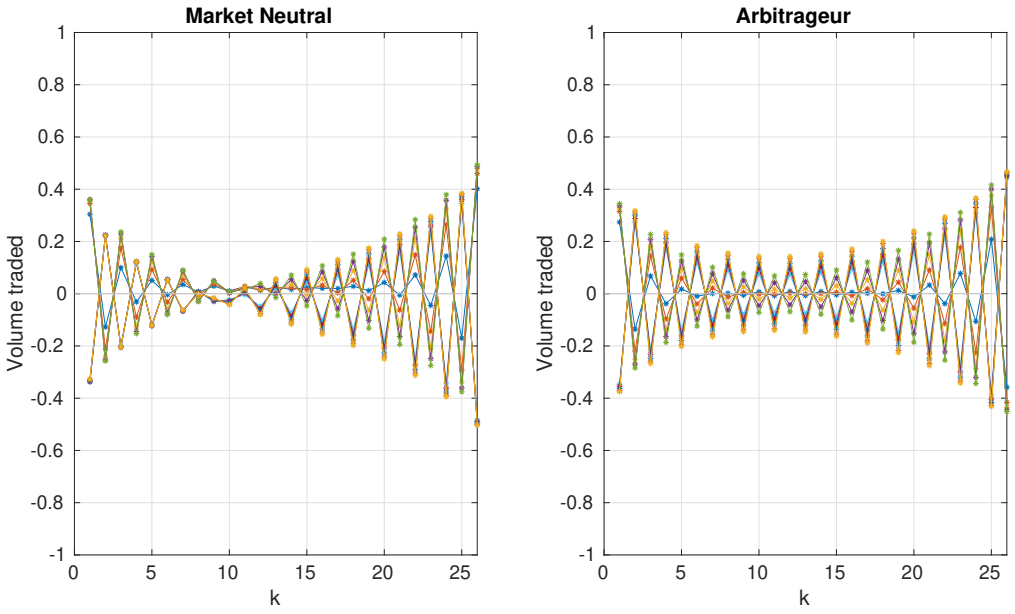


Fig. 9.3 Nash equilibrium when $\theta = 1.5$ with inventories for the Market Neutral $X_0 = (1, \ldots, 1, -1, \ldots, -1)^T \in \mathbb{R}^M$ and for the Arbitrageur $Y_0 = (0, \ldots, 0)^T \in \mathbb{R}^M$, where $M = 2000$. The cross impact matrix is a block matrix with $K = 10$. The figure exhibits the equilibria related to one (the first) asset for each block. The trading time grid is an equidistant time grid with 26 points. Each block has a cross-impact q_i equal to $0.1, 0.2, \ldots, 0.9$ for $i = 1, 2, \ldots, 9$ and 0.95 for the last one.

to infinity at least one of the cluster will increase to infinity, which means that there exists a cluster such that $\lambda_1(Q_i - q\boldsymbol{e_i}\boldsymbol{e_i}^T) \to \infty$ and also the respective eigenvalue of Q goes to infinity. Therefore, we conclude for Theorem 9.4.

For the general case we conclude by contradiction. If $K(M)$ is the number of cluster for a fixed M, and $K(M) \to \infty$ when $M \to \infty$ then the set $\{M_i : i \in \mathbb{N}\}$ is unbounded. Indeed, if $\sup_{i\in\mathbb{N}} M_i = S < \infty$, then the average number of stocks in a cluster is $\frac{\sum_{i=1}^{K(M)} M_i}{K(M)} \leq S$ for all M and this is in contradiction with the assumptions that $\lim_{M\to+\infty} \frac{M}{K(M)} \to +\infty$. So since $\{M_i : i \in \mathbb{N}\}$ is unbounded we conclude that there is no finite value of θ such that it is greater than all the eigenvalues of Q when $M \to \infty$. □

As an example, we consider $K = 10$ equally sized blocks from an universe $M = 2,000$ assets and set $q = 0.05$. With this kind of cross impact matrix, we have K large eigenvalues whose eigenvectors correspond to virtual assets displaying oscillations. The optimal trading strategies for stocks belonging to the same block are the same. Thus in Figure 9.3 we show the Nash equilibrium for the first asset in each of the 10 blocks when the two agents are a Market Neutral and an Arbitrageur. The oscillations are evident, as expected, in all traded assets.

We now study how the critical value θ^* varies when the number of assets increases for different structures of the cross impact matrix and therefore of the liquidity factors.

Comparing different matrix structures is not straightforward since the critical value depends on the values of the matrix elements. To this end we consider the set of symmetric cross impact matrices of M assets having one on the diagonal and fixed sum of the off diagonal elements. More precisely let $h \in \mathbb{R}$, then we introduce for each M the set

$$\mathcal{A}_h^M := \{A \in \mathbb{R}^{M\times M} | A^T = A,\ \sum_{j=1}^{N}\sum_{i>j} a_{ij} = h,\ a_{ii} = 1\},$$

One important element of this set is the cross impact matrix $Q_{1fac} \in \mathbb{R}^{M\times M}$ of a one factor model (see above) with off-diagonal elements equal to $2h/M(M-1)$.

Theorem 9.7. *For a fixed $h \in \mathbb{R}$, let us consider the related one-factor matrix $Q_{1fac} \in \mathcal{A}_h^M$, then*

$$\lambda_1(Q) \geq \lambda_1(Q_{1fac}), \quad \forall Q \in \mathcal{A}_h^M,$$

i.e., among all the matrices with one in the diagonal and constant sum of the off-diagonal terms, the one-factor matrix (i.e., where all the off-diagonal elements are equal) is one of the matrices with the smallest largest eigenvalue.

Proof of Theorem 9.7. The largest eigenvalue of a symmetric $M \times M$ matrix Q can be defined as

$$\lambda_1(Q) = \max_{\boldsymbol{x}\neq 0} \frac{\boldsymbol{x}^T Q \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}}.$$

If we consider the vector $\boldsymbol{e} = (1, 1, ..., 1)^T$, we have the lower bound

$$\lambda_1(Q) \geq \frac{\boldsymbol{e}^T Q \boldsymbol{e}}{\boldsymbol{e}^T \boldsymbol{e}} = \frac{\sum_{i,j} q_{ij}}{M}.$$

The largest eigenvalue of a generic matrix $Q \in \mathcal{A}_h^M$ is then bounded by

$$\lambda_1(Q) \geq 1 + \frac{2h}{M}.$$

But the one-factor matrix $Q_{1fac} = (1-q)I_M + q\boldsymbol{e}\boldsymbol{e}^T$, with $q = \frac{2h}{M(M-1)}$, belongs to $\mathcal{A}_h^M$ and has

$$\lambda_1(Q_{1fac}) = 1 + (M-1)\frac{2h}{M(M-1)} = 1 + \frac{2h}{M},$$

i.e., the lower bound for the max eigenvalue of matrices in $\mathcal{A}_h^M$. Therefore, $\forall Q \in \mathcal{A}^h$ it holds that

$$\lambda_1(Q) \geq \lambda_1(Q_{1fac}).$$

□

Moreover, we note that the previous is not a strict inequality. Indeed, both a diagonal

block matrix, with identical blocks, and the one-factor matrix have the same maximum eigenvalue. Let consider the block diagonal matrix with K identical clusters

$$Q := \begin{bmatrix} Q(\rho) & 0 & \cdots & 0 \\ 0 & Q(\rho) & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & Q(\rho) \end{bmatrix} \in \mathbb{R}^{M\times M},$$

where $Q(\rho) \in \mathbb{R}^{M_c}$ is a one-factor matrix and $M_c \cdot K = M$. We observe that $Q \in \mathcal{A}_h^M$ if and only if $\rho = \frac{2h}{(M_c-1)M}$, therefore

$$\lambda_1(Q) = 1 + (M_c - 1)\rho = 1 + \frac{2h}{M}.$$

This theorem implies that among all the cross impact matrices belonging to $\mathcal{A}_h^M$, the one factor case is among the most stable cross-impact matrices. For example, it is direct to construct an example of a block diagonal cross impact matrix with non-zero off block elements (i.e., similar to what observed empirically) and to prove that its critical θ^* is larger than the critical value for the one factor matrix having the same value h of total cross-impact.

In Appendix C we analyze the case of two risk-neutrals agents trading one asset with different trading skills.

9.2 Market stability in multi-agent and multi-asset market impact games

We now study how the stability of the market depends on the number of agents, J, together with the number of assets, M, risk-aversion parameter γ, and number of trading times N. Specifically, we compute numerically the critical value of θ after which the market is not stable. However, we first observe that to study the stability it is sufficient to analyze the fundamental solutions of each virtual assets.

9.2.1 Characterization of the fundamental solutions

If all agents have the same inventory, i.e., $\boldsymbol{X}_{\cdot,j} = \boldsymbol{Z}\ \forall j$ where $\boldsymbol{Z} \in \mathbb{R}^M$ is a fixed inventory, then also the virtual inventories are all equal, since $\boldsymbol{X}^P_{\cdot,j} = V^T\boldsymbol{Z} \equiv \boldsymbol{Z}^P\ \forall j$. Then, $\overline{X}^P_{i,\cdot} = \frac{1}{J}\sum_{j=1}^J X^P_{i,j} = Z^P_i$ and by Eq. (7.5) the solution for all agent j in virtual asset i is given by $\Xi^{*,P}_{i,j} = Z^P_{i,j}\boldsymbol{v}_i$. So, let $V = \left[\boldsymbol{\nu}_1|\boldsymbol{\nu}_2|\cdots|\boldsymbol{\nu}_M\right]$ the matrix of eigenvectors of Q, which we may assume to be normalized, $\boldsymbol{\nu}_i^T\boldsymbol{\nu}_i = 1$, if $\boldsymbol{X}_{\cdot,j} = \boldsymbol{\nu}_m\ \forall j$ then the

optimal schedule on the virtual assets is given

$$\Xi^{*,P}_{i,j} = \begin{cases} \boldsymbol{v}_m, & i = m \\ \mathbf{0}, & \forall\, i \neq m \end{cases}, \quad \forall j$$

since $\boldsymbol{X}^P_{\cdot,j} = V^T \boldsymbol{X}_{\cdot,j}$ has 1 in the m-th position and zero otherwise, so $\Xi^*_{\cdot,j} = V \cdot \Xi^{*,P}_{\cdot,j} = \boldsymbol{\nu}_m \otimes \boldsymbol{v}_m,\ \forall j$, which means that the strategies for all traders is fully characterized by the fundamental solution $\boldsymbol{v}_m$.

If $\overline{X}_{i,\cdot} = 0,\ \forall i$ then $\overline{X}^P_{i,\cdot} = 0$ and by Eq. (7.5) the solution for each agent j is given by $\Xi^{*,P}_{i,j} = X^P_{i,j} \boldsymbol{w}_i,\ i = 1, 2, \ldots, M..$ Thus, as for the previous case, if the inventory of the j-th trader $\boldsymbol{X}_{\cdot,j} = \boldsymbol{\nu}_m$ (and if $\overline{X}_{i,\cdot} = 0$ for all i), then his/her optimal schedule on the virtual assets is given by

$$\Xi^{*,P}_{i,j} = \begin{cases} \boldsymbol{w}_m, & i = m \\ \mathbf{0}, & \forall\, i \neq m \end{cases},$$

so that $\Xi^*_{\cdot,j} = V \cdot \Xi^{*,P}_{\cdot,j} = \boldsymbol{\nu}_m \otimes \boldsymbol{w}_m,$.

We summarize the previous results as follows:

a) If all agents have the same inventories, i.e. $\boldsymbol{X}_{\cdot,j} = \boldsymbol{\nu}_m\ \forall j$, then the Nash equilibrium for j is proportional to $\boldsymbol{v}_m$, i.e, $\Xi^*_{\cdot,j} = \boldsymbol{\nu}_m \otimes \boldsymbol{v}_m$.

b) If $\overline{X}_{i,\cdot} = 0,\ \forall i$ and $\boldsymbol{X}_{\cdot,j} = \boldsymbol{\nu}_m$, then the Nash equilibrium for j is proportional to $\boldsymbol{w}_m$, i.e, $\Xi^*_{\cdot,j} = \boldsymbol{\nu}_m \otimes \boldsymbol{w}_m$.

We observe that, respectively, if $\boldsymbol{v}_m$, or $\boldsymbol{w}_m$, exhibits spurious oscillations also $\Xi^*_{\cdot,j}$ is affected by these oscillations, respectively. We recall that market is unstable if a particular initial inventory leads to optimal trading strategies with spurious oscillations. So we can restrict the stability analysis on the fundamental solutions among all assets.

9.2.2 Numerical analysis of stability

From the results of Section 9.1 we known that market stability is affected by the cross-impact structure in a market with two risk-neutral agents. Thus, in this section we want to study how the number of agents J, the risk-averse parameter γ together with the number of assets M might affect the market stability in the multi-agent and multi-asset case. We also examine the role of number of trading step N, even if we expect to have no role in stability, as also observed by Luo and Schied (2020) for the one asset case. In particular, we compute numerically θ^* such that when $\theta < \theta^*$ the market is unstable. As observed in Section 9.2.1 it is sufficient to examine the oscillations of the fundamental solutions on the virtual assets.

We consider the following setting with J risk-averse agents, where γ is the risk-averse parameter, and M assets:

- The time grid is equidistant $\mathbb{T}_N = \{\frac{kT}{N} | k = 0, 1, \ldots, N\}$, where $T = 1$ and $N \in \mathbb{N}$;
- The decay kernel is exponential, $G(t) = e^{-t}$;
- The cross-impact matrix is a one factor matrix, $Q = (1-q)I_M + q\boldsymbol{e}\boldsymbol{e}^T$, where $q = 1/2$;
- $\boldsymbol{S}_t^0$ follows a Bachelier model where the covariance matrix is equal to Q.

The study of Luo and Schied (2020) points out a conjecture on θ^* in the one-asset case, where it comes up that

$$\sup_{N,\gamma} \theta^*(1, J, N, \gamma) = G(0) \cdot \frac{J-1}{4},$$

therefore, given the results of Section 9, our conjecture is that

$$\sup_{N,\gamma} \theta^*(M, J, N, \gamma) = G(0) \cdot \frac{(J-1)\lambda_{max}}{4}, \tag{9.2}$$

where λ_{max} is the maximum eigenvalue of Q. We recall that in the above setting, $\lambda_{max} = 1 + \frac{M-1}{2}$ and $G(0) = 1$. Thus, in the first analysis we set $N = 300$, $\gamma = 10$ and we compute θ^* as a function of M and J. Figure 9.4 exhibits the corresponding level curves. It is worth noticing that the relation between J and M is very close to that of Equation 9.2. Indeed, the average relative discrepancy on θ^* is of the order of 10^{-3}. Finally, we examine how θ^* depends on N and γ for fixed M and J, which are $M = J = 11$, see Figure 9.5 which illustrates the related surface[3]. Overall, the numerical results suggests that for fixed M and J the relation (9.2) holds when N is not too small, since for the chosen parameter Eq. (9.2) predicts $\theta^* = 15$.

Therefore, under generic assumptions, when either the number of agents J or of assets M increase, market turns out to be unstable unless the transaction costs parameter θ increases appropriately. This result depends on the assumption that impact strength does not depend on the number J of agents. In Remark 7.3 we observed that there are theoretical arguments suggesting that the kernel depends on J as $G^{sc}(t) := J^{-\beta} \cdot G(t)$, where $\beta \geq 0$ is the scaling parameter. The question is how the critical value θ^* depends on J in this case. We observe that for the existence and uniqueness of Nash equilibria Theorem 7.8 still holds, i.e., $G^{sc}(t)$ is a scaled version of $G(t)$ and so it preserves the same property of strictly positive definiteness of $G(t)$. Furthermore, we observe that

[3]We also compute the same surface for $M = J = 3$ and $M = J = 5$, and we obtain similar results, available upon request.

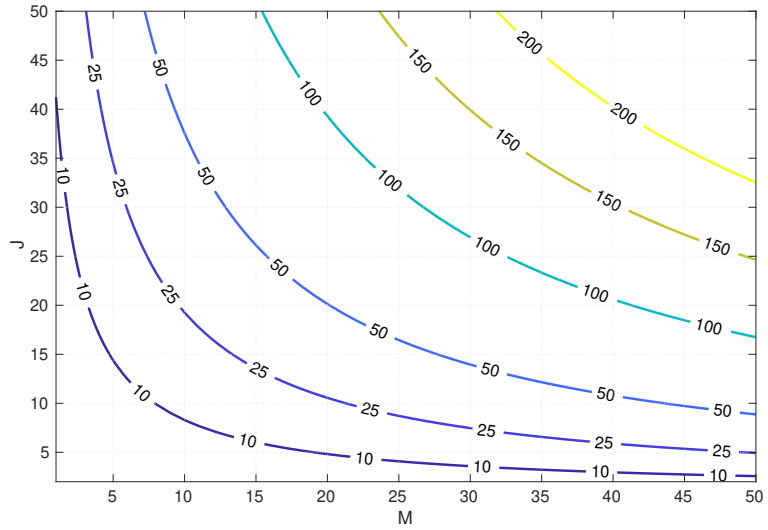


Fig. 9.4 Level curves of θ^* in function of M and J for fixed $N = 300$ and $\gamma = 10$. The bottom left corner, corresponding to $M = 1$ and $J = 2$, is the case of Schied and Zhang (2018).

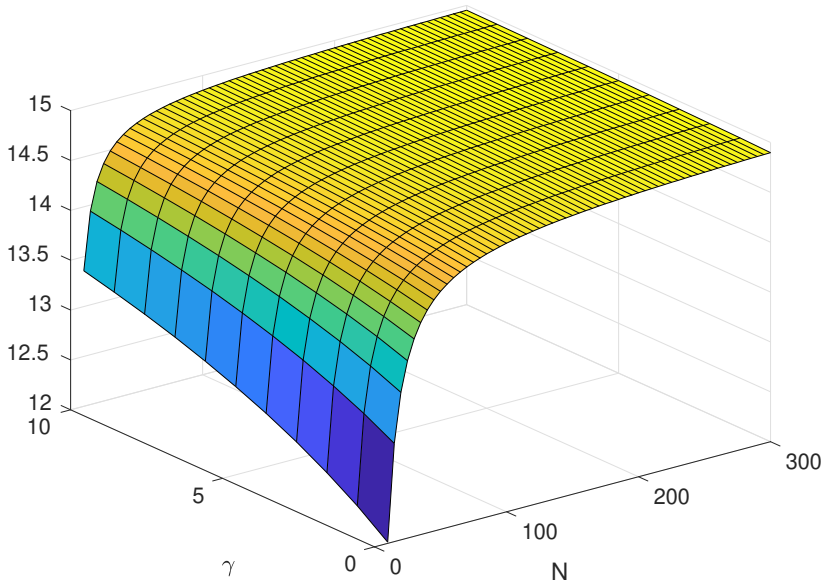


Fig. 9.5 Surface plot of θ^* in function of N and γ for fixed $M = J = 11$.

all the previous analyses are performed with $\beta = 0$. Therefore, according to relation (9.2), if we introduce this scaling parameter β, we expect that when M and J are fixed $\sup_{N,\gamma} \theta^*(M, J, N, \gamma)$ decreases with β, since $G^{sc}(0) = J^{-\beta} G(0)$. It is therefore expected that the critical transaction cost level is a decreasing function of β.

We numerically compute the value of θ^* as a function of β and J by fixing $M = 50$, $N = 50$, and $\gamma = 10$, and we plot in Figure 9.6 the contour plot. As expected, for fixed J the critical transaction cost level is a decreasing function of β. Moreover, replacing G

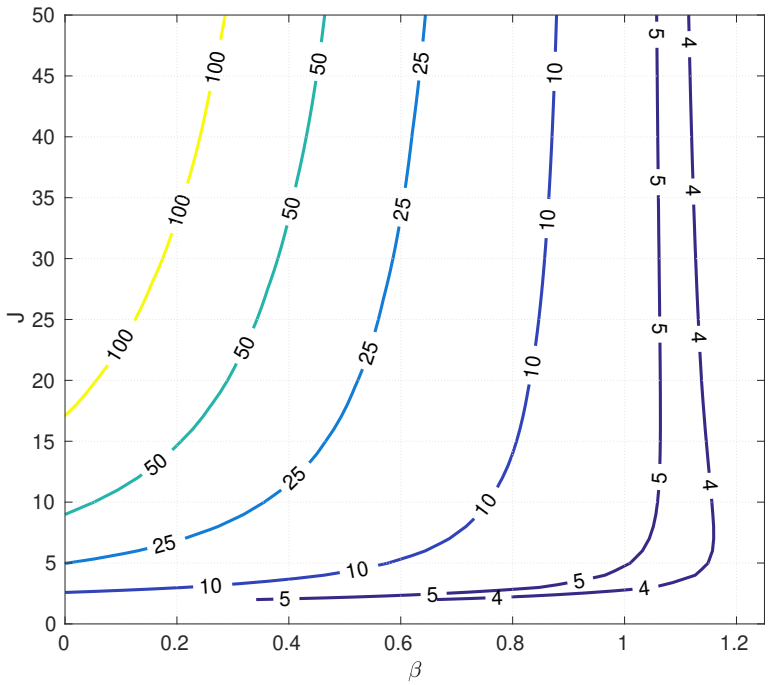


Fig. 9.6 Numerical estimates of level curves of θ^* as a function of β and J for fixed $M = 50$, $N = 50$, and $\gamma = 10$. The level curves computed according to relation (9.2) are very similar, with a relative discrepancy of the order of $2 \cdot 10^{-2}$.

with its scaled version G^{sc}, we compute the average relative error with respect to (9.2), which we find to be of the order of $2 \cdot 10^{-2}$. Again, the numerical results do not reject the conjecture (9.2). Therefore, according to relation (9.2), when all other parameters are fixed the critical value is driven by the ratio

$$\frac{J-1}{J^{\beta}},$$

and we have three possible scenarios.

- $\beta > 1$. Market is more prone to stability. An increasing in competition shall act as a stabilization effect which will reduce the critical value,
- $0 \leq \beta < 1$. Market is more prone to instability, since the critical value increases with J.
- $\beta = 1$. This is an uncertain scenario, since we do not known if an increasing in the number of traders may affect market stability.

We have mentioned in Remark 7.3 that some theoretical arguments suggest $\beta = 1/2$, while empirical studies provide evidence that $\beta > 0$. Fig. 9.6 shows that the critical value θ^* strongly depends on the scaling exponent β, thus its estimation is determinant for assessing the stability properties of markets.

9.2.3 Possible policy recommendations

We conclude by briefly presenting some policy recommendations we draw from the model when the objective is to avoid the occurrence of instabilities. The conjecture above indicates that the critical transaction cost level θ below which instabilities are present grows with the impact coefficient $G(0)$ (or its scaled version), the number of traders J, and the largest eigenvalue λ_{max} of the cross impact matrix. The latter quantity is typically an increasing function of the number of assets M, for example when the cross-impact is described by a one-factor matrix. Thus, to ensure stability, transaction cost parameter θ should be set taking into account the above variables, and be increased or decreased when they significantly change[4].

Clearly, an increase of the transaction costs might discourage trading activity, therefore decreasing overall market participation and possibly price discovery. For example, in the one period multi-agent Kyle model of Bagnoli et al. (2001) the mean square deviation of the market price from the fundamental value goes to zero with the number of agents as $(J+1)^{-1}$. Thus regulators should fix transaction costs by balancing the contrasting objectives of increasing traders participation/price discovery and stabilizing markets.

An important aspect to consider in this trade-off is the way in which market impact of a single agent depends on the number of agents, i.e., what we modeled with the scaled impact G^{sc}, since β affects significantly θ^*. Despite some theoretical and empirical results are available (see Remark 7.3), this is still an open issue, which is certainly worth of investigation. A policy regulator may decide to increase or reduce transaction costs to stabilize market depending on the scaling parameter β.

[4]In principle, regulators could also act on $G(0)$ by implementing measures making the market more liquid to individual trades, for example modifying the cost of limit orders.

Appendix C

Instability and heterogeneity of agents' trading skills

While in the Schied and Zhang framework the two investors are identical, real markets are characterized by huge heterogeneity in trading skills. For example, some agents (e.g. HFTs) are much faster than others, some agents use more sophisticated trading strategies and have smaller trading costs, etc. Is this traders' heterogeneity beneficial to market stability? Do HFTs destabilize markets? In this Appendix we study the case of two risk-neutrals agents trading one asset, but, differently from Schied and Zhang (2018), we assume that the two agents are different in their trading skills. We consider separately three sources of heterogeneity: different permanent impact, different temporary impact (or trading fees), and different trading speed. If not specified we identify traders with their respective inventory and we denote the first and second trader as agent X and Y, respectively.

In the first setting, we study the case when the two agents affect with a different impact the price process. This might correspond to more sophisticated traders who are able to trade more efficiently, for example using better algorithms for posting orders in the market. There is an alternative interpretation, more in line with traditional market microstructure. Permanent impact is considered a measure of the informativeness of trades, thus an agent with a very small permanent impact might be interpreted as a noise trader. In our model the two interpretations are indistinguishable. For analytical convenience, we assume the kernel $G(t)$ is the same for the two agents and we introduce two scaling parameters β_i $(i = 1, 2)$ such that the price process dynamic affected by trading becomes[1]

$$S_t^{\Xi} = S_t^0 - \sum_{k=0}^{N} G(t - t_k)(J^{-\beta_1}\xi_{1,k} + J^{-\beta_2}\xi_{2,k}), \tag{C.1}$$

where $J = 2$. We observe that when two agents have different β's this does not compro-

[1]Note that this is equivalent to saying that the $G(t)$ of the two agents are proportional one to each other, i.e. the two agents differ in the value of $G(0)$.

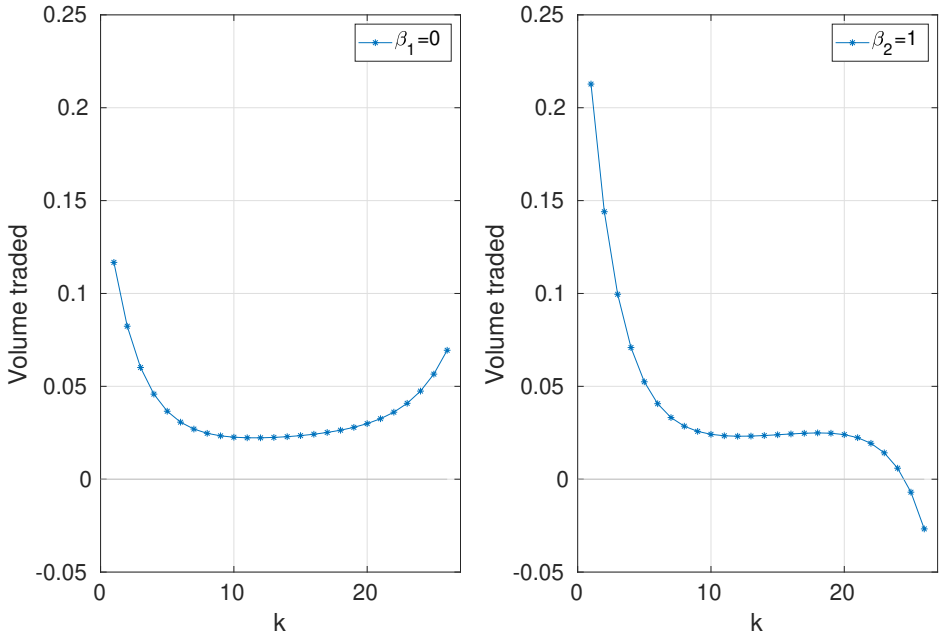


Fig. C.1 Nash equilibrium for two Directional sellers when $\theta = 1.5$. The left (right) figure is referred to the agent with $\beta_1 = 0$ ($\beta_2 = 1$). The trading time grid is an equidistant grid with 26 points and the decay kernel is $G(t) = \exp(-t)$.

mise the market stability. Indeed, the previous equation can be rewritten as

$$S_t^{\Xi} = S_t^0 - \sum_{k=0}^{N} G(t - t_k)(\widetilde{\xi}_{1,k} + \widetilde{\xi}_{2,k}), \tag{C.2}$$

where $\widetilde{\xi}_{1,k}$, $\widetilde{\xi}_{2,k}$ are the strategies associated with the transformed inventories $\widetilde{X}_1 = J^{-\beta_1} X_1$, $\widetilde{X}_2 = J^{-\beta_2} X_2$. Therefore, the Nash equilibrium of Eq. C.1 is related to the equilibrium of Eq. C.2, i.e., the are equal up to constants, and in particular the stability is not compromised. We just mention that if, for instance, $\beta_1 = 0$ and $\beta_2 = 1$, i.e. the impact of the second agent is half the one of the first agent, and both agents are Directional sellers, the sophisticated agent Y has the advantage as if he/she should trade half of his volume. Figure C.1 exhibits the Nash equilibrium for both traders. As expected, the optimal trading pattern is different for the two traders. More interestingly, the trading profile of the agent with small impact resembles the one of an Arbitrageur (see for comparison Fig. 8.1). Thus, from the point of view of the trading profile, an Arbitrageur with zero inventory or a Fundamental trader with small impact behaves in a similar way at equilibrium. Remember that the latter can be seen either as a skilled trader or as a noise trader. Moreover, the expected cost for the $\beta = 0$ and $\beta = 1$ sellers are equal to 0.6521 and 0.2582, respectively. When both agents have $\beta = 0$, the costs are equal to 0.7975, thus the reduction of impact of one agent is indeed beneficial to both traders.

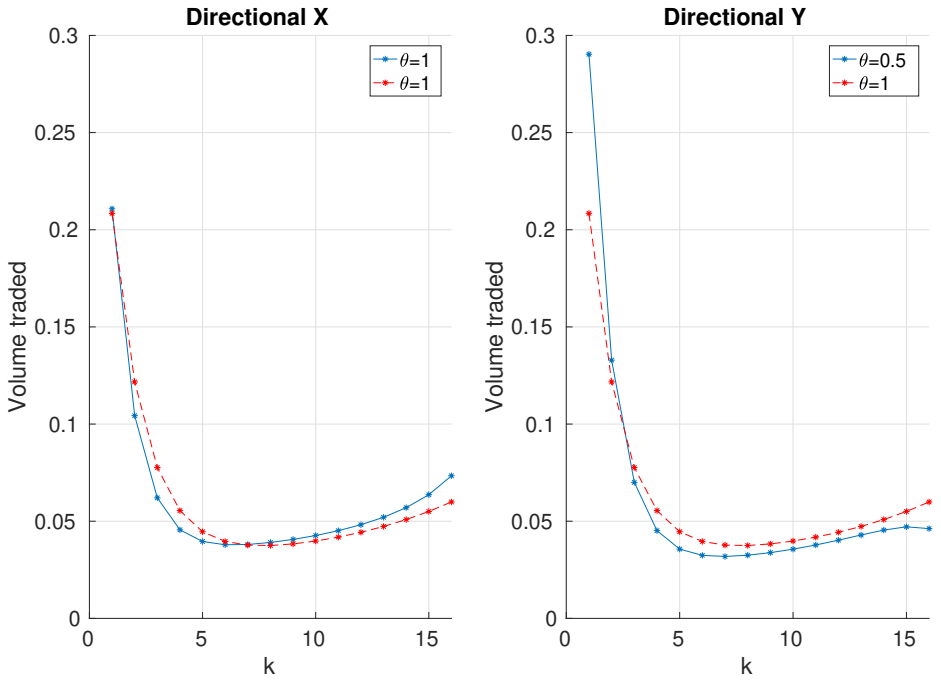


Fig. C.2 Red lines are the Nash equilibrium when both traders have $\theta = 1$. Blue lines are the solution when the first trader has $\theta = 1$ and the second has $\theta = 0.5$. The trading time grid is an equidistant grid with 16 points and the decay kernel is $G(t) = \exp(-t)$.

In the second setting, we consider two agents with different temporary impact parameter θ. This might correspond again to the case where one of the two agents has more sophisticated trading strategy or to a market where different class of traders are allowed to pay different fees. This is indeed quite standard for example in ATS where liquidity providers have lower fees than liquidity takers. In particular, we want to test whether only one agent with θ below threshold is sufficient to destabilize the market or if both agents must have $\theta < \theta^*$. We denote with θ_X and θ_Y the temporary impact parameter for the two agents. We study the Nash equilibrium by solving numerically the Schied and Zhang (2018) game in the setting where both agents are Directional seller with $X_1 = X_2 = 1$. When both agents are above threshold ($\theta^* < \theta_Y < \theta_X = 2\theta_Y$, see Figure C.2), we observe that the more significant difference between the two solutions is in the trading schedule of low θ agent, who trades more at the beginning of the trading session. It is interesting to consider the changes in cost. When $\theta = 1$ for both agents, the individual cost is 0.7970. Suppose now that θ_Y is reduced to 0.5 (for example because a new type of agent enters the market or because the fee schedule changes for some class of traders). As expected, the expected cost of agent Y is reduced (0.7317), but there is a negative effect on the other agent X who sees her cost soar to 0.8286. Thus, the decrease of θ for a class of agents impacts negatively the other market participants.

When we examine the case when one of the agents has a θ_Y smaller than the critical value $\theta^* = 0.25$, we find numerical evidences that the equilibrium becomes unstable also

Table C.1 Expected costs of the two agents X and Y as a function of the trading speed p (probability of arriving first on the market) of Y. When $p = 1$ the agent Y places orders always before of X.

$\mathbb{E}[C_T(\boldsymbol{\xi}_1\|\boldsymbol{\xi}_2)]$	$\mathbb{E}[C_T(\boldsymbol{\xi}_1\|\boldsymbol{\xi}_2)]$	p
0.7970	0.7970	1/2
0.8173	0.7765	2/3
0.8274	0.7662	3/4
0.8333	0.7601	4/5
0.8568	0.7360	1

for the Directional agent for which $\theta_X > \theta^*$. This suggests that in a market with few sophisticated traders with small temporary impact (HFTs) or agents allowed to pay lower fees, the instability comes up for all the agents in the market, regardless for their own θs. In other words, market instability is driven by the agent with the smallest θ and having even few sophisticated traders (in terms of costs or fees) destabilizes the market.

The third and last setting assumes that the two traders are different because one of the two is faster than the other. A fast trader could represent, for example, an HFT. In market impact games one has to decides who arrives first in the market at each time step, because the laggard is going to pay the impact of the trade of the leader. In the standard setting, this is decided by a fair Bernoulli trial, which means that on average each trader is the leader half of the times. We modify this and for each trading time t_k the Bernoulli game ε_k deciding the trading priority, is no longer fair, i.e.,

$$\varepsilon_k = \begin{cases} 1, & \text{with probability } p \\ 0, & \text{with probability } 1-p \end{cases} \qquad p \in [0,1]$$

where $\varepsilon_k = 1$ means that Y places the k-th order before X, so that X has to take into account the impact of Y in her cost. Let us denote with $\boldsymbol{\xi}_1$ ($\boldsymbol{\xi}_2$) the Nash equilibrium for X (Y), then we observe that $\mathbb{E}[C_T(\boldsymbol{\xi}_1|\boldsymbol{\xi}_2)] = \mathbb{E}[\frac{1}{2}\boldsymbol{\xi}_1^T\Gamma_\theta\boldsymbol{\xi}_1 + \boldsymbol{\xi}_1^T\Gamma_p\boldsymbol{\xi}_2]$ and $\mathbb{E}[C_T(\boldsymbol{\xi}_2|\boldsymbol{\xi}_1)] = \mathbb{E}[\frac{1}{2}\boldsymbol{\xi}_2^T\Gamma_\theta\boldsymbol{\xi}_2 + \boldsymbol{\xi}_2^T\Gamma_{1-p}\boldsymbol{\xi}_1]$, since

$$\sum_{k=0}^{N}\xi_{1,k}\left(p\xi_{2,k} + \sum_{m=0}^{k-1}\xi_{2,m}G(t_k - t_m)\right) = \boldsymbol{\xi}_1^T\Gamma_p\boldsymbol{\xi}_2 \quad \text{where } (\Gamma_p)_{ij} = \begin{cases} \widetilde{\Gamma}_{ij}, & i \neq j \\ p\cdot G(0), & i = j \end{cases}.$$

Figure C.3 exhibits the Nash equilibrium of the two traders for several values of p and Table C.1 reports the corresponding expected costs. We observe that when p increases, the optimal solution for the HFT Y is to liquidate slightly faster in the first period and then sell the remaining part of the inventory with a lower intensity, with respect to the solution with $p = 0.5$, at the end of the trading session. For the slow trader X

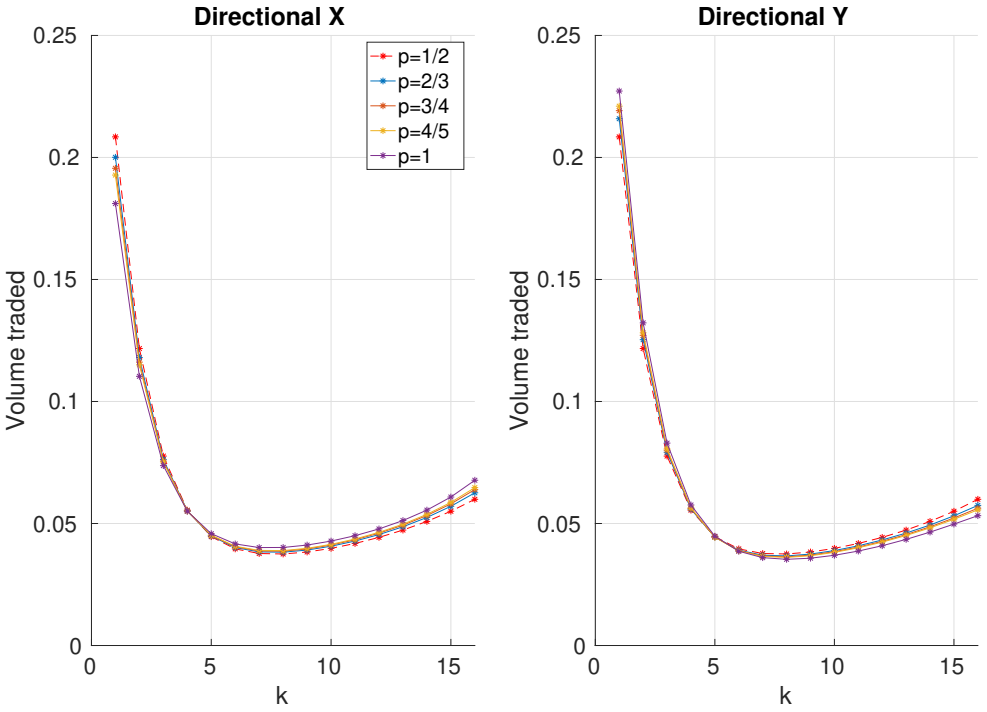


Fig. C.3 Nash equilibrium in function of the trading speed p (probability of arriving first on the market) of Y. The red dotted lines are the Nash equilibrium when both traders have the same trading speed, i.e., $p = 0.5$. The time grid has 16 points, $\theta = 1$ and $G(t) = \exp(-t)$.

the behavior is exactly opposite to the one of the fast trader. However the differences between Nash equilibria for different p are quite small, while the greatest benefit for the HFT is the smaller trading costs (see Table C.1). Moreover, ask whether the presence of a fast trader modifies the critical value of θ when the instability starts. Figure C.4 shows how θ^* varies as a function of the trading speed p for a games between two Directional sellers[2]. Since we compute the critical value using a numerical method there are some small oscillations. We observe a small but significant trend of θ^* as a function of p. This result indicates that the presence of an HFT makes the market more prone to oscillations and instabilities. However the effect is relatively small compared to the other possible sources of instabilities studied in the previous sections.

Finally, we analyze the game between a Directional seller X and an Arbitrageur Y, when $p \in [0, 1]$. We recall that when p is close to 1 the Arbitrageur Y places orders almost always before of the Directional. Figure C.5 exhibits how the shape of the equilibrium varies as a function of p. In particular, the solution of the Arbitrageur is very weakly affected by p, while for the Directional we observe that when p is close to 0 the optimal schedule converges to a U-shape form, putting more order at the beginning at the session

[2]We also repeat the same experiment in the presence of an Arbitrageur against a Directional and we find analogous results.

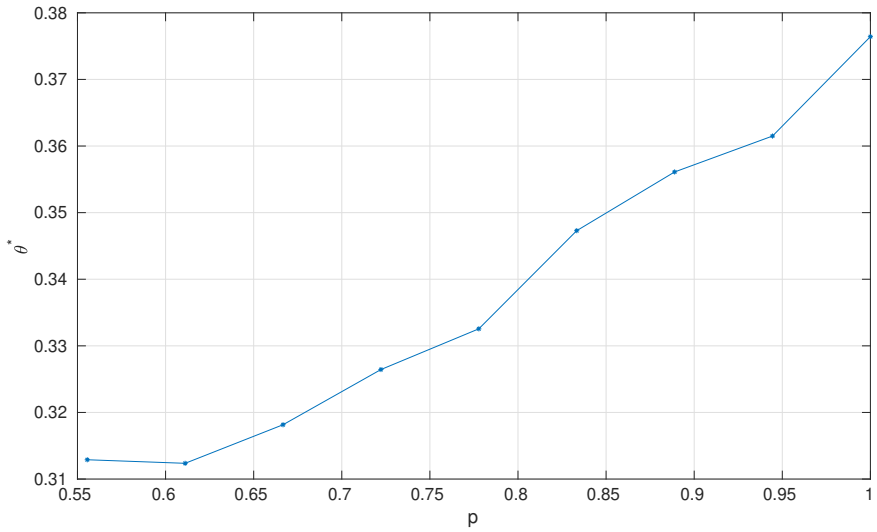


Fig. C.4 Numerical estimates of the critical value θ^* as a function of the trading speed p (probability of arriving first on the market) of the HFT. We set $N = 6$ for the time grid and $G(t) = \exp(-t)$. Both agents are Directional sellers.

compared to the standard solution with $p = 0.5$. This can be motivated by the fact that when X always places orders before Y basically there is no game[3] for the Directional and the optimal solution is given by the classical U-shape of the TIM. However, it seems to be suggestive that the U-shape is the optimal way to exploit the impact of the Arbitrageur when the Directional is an ultra-fast trader. Vice versa, when p is close to 1, the solution of the Directional tends to concentrate more order at the end of the trading session similarly to the standard solution with $p = 0.5$. So, we may conclude that when the Arbitrageur is very fast with respect to the Directional, the optimal schedule is provided by an asymmetric solution where the seller tends to concentrate more orders when the Arbitrageur puts buy orders.

[3]Since Y is an Arbitrageur there is no competition to liquidate the asset.

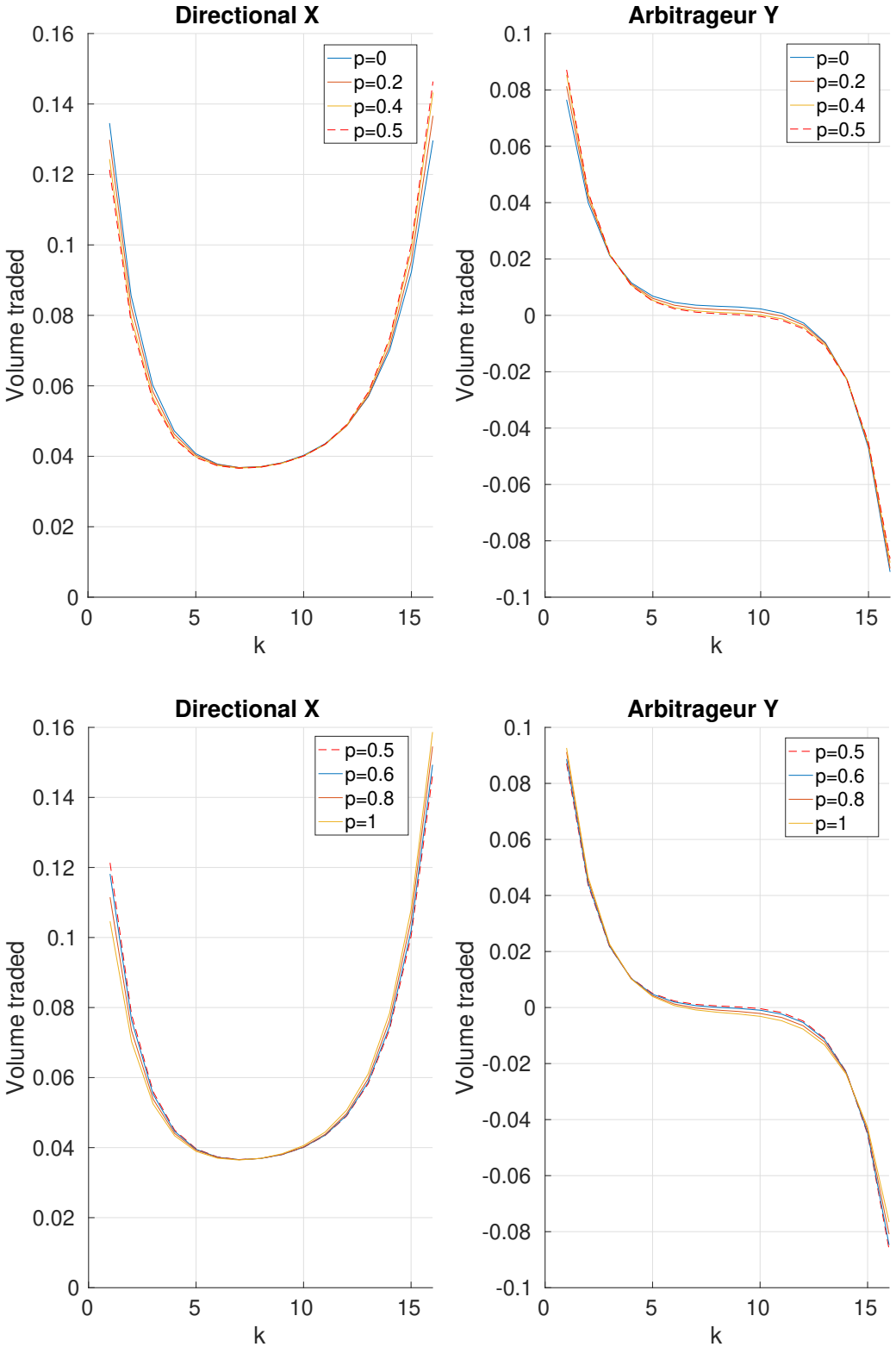


Fig. C.5 Nash equilibrium as a function of the trading rate p of the Arbitrageur Y. The red dotted lines are the Nash equilibrium when both traders have the same trading rate, i.e., $p = 0.5$. The time grid has 16 points, $\theta = 1$ and $G(t) = \exp(-t)$.

CONCLUSION

In this thesis we started by looking for possible mispricing relations to build profitable investment strategies and we finally analyzed how the (microscopic) interactions among traders may affect trading conditions and market instability. We summarize in the following points the main contributions.

- In Chapter 2, we have considered generalizations of the recently proposed ICA identification approach for SVAR models in two directions: the singular and noisy case. We have shown how to combine the collapsing procedure, developed in the context of a state space model, together with the ICA technique, in a new identification approach that we term Collapsing-ICA. We have discussed in details under which set of assumptions the proposed methodology remains feasible and the consistency of the procedure as T goes to infinity. In particular, we have shown that when the VAR is singular we can identify the structural shocks from the reduced ones thus preserving the fundamentalness property. Moreover, when the noise is homoskedastic and the system amplification is homogeneous we can consistently estimate the mixing matrix Z and so we may conduct the IRFs analysis directly on the structural shocks. In the singular case, standard ICA identification techniques would recover the IRFs for all the selected variables. However, this could be misleading in this context as it implicitly assumes more structural shocks than the ones actually driving the system. On the other hand, the C-ICA procedure is well adapted to singular cases allowing to use all the information contained in the observed variables to consistently identifying the lower dimensional system of the structural shocks. The empirical analysis is in agreement with the theoretical and simulated results. We have compared the IRFs obtained with the identification scheme of Gouriéroux et al. (2017) on a system with 3 variables, with that of C-ICA on a system of 4 variables, including the output gap, which was instead discarded by Gouriéroux et al. (2017) in favor of the unemployment gap. The IRFs analysis and the variance decomposition confirm that C-ICA correctly merges the information contained in all the variables thus allowing a more precise and coherent identification of the low dimensional system of structural shocks driving the economy.

- In Chapter 3, we have proposed a framework that can be viewed as a generalization of the traditional DCF model of firm valuation, in which the point estimate of the

traditional approach is replaced with an estimated probability distribution of fair values. In this way one can derive both an estimate of the fair value of a company and a measure of the degree of uncertainty associated with it. The proposed valuation framework is named Stochastic Discount Cash Flow method, SDCF, which is rooted in fundamental analysis and based on an econometric forecasting model of future firm's cash flow. We have stressed that our approach is invariant to the considered valuation model.

- In Chapter 4, we have shown the importance of a distributional approach to valuation through different exercises to investigate the misvaluation effect. We have shown that a simple volatility-adjusted misvaluation indicator, derived from the estimated fair value distribution, possess predictive power with respect to stocks future returns. Moreover, by longing undervalued stocks and shortening overvalued stocks, we are able to build a misvaluation factor, the long-short valuation LSV factor, that captures novel information not accounted for in previously explored factor models and posses a significant explanatory power of realized abnormal returns of both portfolios and individual stocks. In this spirit, our analysis is similar to those performed in Hirshleifer and Jiang (2010) and Chang et al. (2013), but our approach is different. In Hirshleifer and Jiang (2010), the authors introduce a misvaluation factor using the special market operations (e.g., repurchase, new issue of equity and debt) the company underwent in the previous two years. In Chang et al. (2013) the company misvaluation is captured by the residual of a sector-wise regression of company's past returns on a set of market factors and a few key firm-specific financial indicators. In both cases, the misvaluation indicator is strictly linked to firm market dynamics and emerges from the comparison of the relative performances, in the long or short term, of different stocks. Conversely, our indicator is based on the comparison of firm's prevailing market prices and the fundamental value estimated starting from balance sheet data, through a careful estimation of the company's future operating performances. The different nature of the misvaluation measure is apparent in their weak correlation and in the fact that our factor loadings are more relevant for the firm level analysis.

- In Chapter 5, we have proposed two recommendation systems based on the comparison of observed market prices with the fair value distributions obtained through the introduced SDCF method. The *Single-Stock Quantile* system derives recommendations for each company, considering only how far the price of the stock is from the median of the computed fair value distribution. The *Cross-Sectional Quantile* system builds a mispricing indicator for each company and then derives recommendations by comparing the indicators across all companies. While the former method fully uses all information available from the fair value distribution, the

latter is more robust with respect to possible fair value estimation biases and also guarantees a constant number of stocks in each recommendation bucket. For each recommendation system, we have built buy and sell side portfolios and we have estimated the abnormal returns, both gross and net trading costs, earned from diverse investment strategies. The *Buy Sides* provides a significant average annual abnormal gross return of about 6% percent, after controlling for market risk, size, book-to-market, and price momentum effects, which doubles the market abnormal gross return (of Our Universe), which is about 3%. Contrary to the portfolios based on analysts' stock recommendations (i.e., the I/B/E/S recommendation system), our investment strategies (portfolios) are always consistent, as buying stocks with a more favorable recommendation invariably earns a greater annualized log-return than buying stocks with less favorable recommendation. Finally, we have shown how to use the uncertainty provided by analysts and thus how to improve their recommendations, by relying on the CSQ methodology, which exhibits the robustness of the cross-sectional approach to issue consistent recommendation system.

Overall, the above results confirm the importance of assessing the degree of uncertainty associated with the valuation of a company. Nonetheless, the present study can be extended in several directions. The univariate models we have adopted in describing the firm's log-revenues dynamics can be replaced by multivariate time-series models, possibly exploiting the cross-sectional information available when explicitly considering the temporal dynamics of different balance-sheet variables. Another relatively straightforward application of the SDCF methodology is comparing the obtained fair value distribution with the company's fair value implied by the price distribution of call/put options, e.g., using the principle of maximum entropy of Buchen and Kelly (1996). This comparison could shed new light on the process by which the temporary misvaluation that our indicator seems able to capture is progressively eliminated by market price adjustments.

Then, we have used market impact games to investigate several potential determinants of market instabilities driven by finite liquidity and simultaneous trade execution of more agents. Specifically, we have extended the results of Schied and Zhang (2018) and Luo and Schied (2020) in several directions.

- First, in Chapter 7, we have considered a multi-asset market where we have introduced the cross-impact effect among assets. We have found the Nash equilibrium, we have analyzed the optimal trading strategies provided by the equilibrium, see Chapter 8, and we have studied the impact of transaction costs on liquidation strategies.

- In Chapter 9, we have studied the stability of the market when the number of assets

increases and we have found that for most realistic cross-impact structures the market is intrinsically unstable. Even if asymptotically the instability arises in all cases, we have found that when the structure of the cross-impact matrix is complex, for example it has a block or multi-factor structure, the instability transition occurs for higher values of the impact parameter. Thus, all else being equal, the temporary impact (or the transaction fees) must be larger in order to observe stability. Finally, we have numerically analyzed market stability in the general model with J risk-averse agents trading M assets. Our results are in agreement with the study of Luo and Schied (2020) and we have found clear evidence that more competition in the market compromises its stability together with an increasing in its complexity (in terms of cross-impact structure). However, when the impact of single agents is scaled by an appropriate parameter, the instability seems to be attenuated, thus leaving an opportunity to policy maker to preserve stability.

An interesting extension of the market impact games framework could be the generalization in a derivative contracts market model, where agents trade derivative contracts together with the related underlying assets. This generalization could also lead to an extension of the proposed stability analysis, where we may analyze the impact of derivative prices on market stability. In particular, we may investigate how oscillations in derivative prices could trigger instabilities in the underlying asset dynamics. However, modeling market impact on derivatives is not straightforward and it can be challenging, even if we consider the one-agent case. A recent contribution in this direction is provided by Tomas et al. (2021), where a Kyle cross-impact model on derivatives is studied, and it can be considered as a first starting point.

REFERENCES

L. Alessi, M. Barigozzi, and M. Capasso. Improved penalization for determining the number of factors in approximate factor models. *Statistics & Probability Letters*, 80 (23-24):1806–1813, 2010.

A. Alfonsi, A. Schied, and A. Slynko. Order book resilience, price manipulation, and the positive portfolio problem. *SIAM Journal on Financial Mathematics*, 3(1):511–533, 2012.

A. Alfonsi, F. Klöck, and A. Schied. Multivariate transient price impact and matrix-valued positive definite functions. *Mathematics of Operations Research*, 41(3):914–934, 2016.

M. Ali, R. El-Haddadeh, T. Eldabi, and E. Mansour. Simulation discounted cash flow valuation for internet companies. *International Journal of Business Information Systems*, 6(1):18–33, 2010.

R. Almgren and N. Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3:5–40, 2001.

B. D. O. Anderson and M. Deistler. Generalized linear dynamic factor models-a structure theory. In *2008 47th IEEE Conference on Decision and Control*, pages 1980–1985. IEEE, 2008a.

B. D. O. Anderson and M. Deistler. Properties of zero-free transfer function matrices. *SICE Journal of Control, Measurement, and System Integration*, 1(4):284–292, 2008b.

D. Ardia and K. Boudt. The peer ratios performance of hedge funds. *Journal of Banking & Finance*, 87:351, 2018.

H. Attias. Independent factor analysis. *Neural Computation*, 11(4):803–851, 1999.

M. Bagnoli, S. Viswanathan, and C. Holden. On the existence of linear equilibria in models of market making. *Mathematical Finance*, 11(1):1–31, 2001.

J. Bai and S. Ng. Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221, 2002.

B. Barber, R. Lehavy, M. McNichols, and B. Trueman. Can investors profit from the prophets? security analyst recommendations and stock returns. *The Journal of Finance*, 56(2):531–563, 2001.

R. Baule and H. Wilke. On the profitability of portfolio strategies based on analyst consensus EPS forecasts. In *28th Australasian Finance and Banking Conference Paper*, 2016.

M. Benzaquen, I. Mastromatteo, Z. Eisler, and J.-P. Bouchaud. Dissecting cross-impact on stock markets: An empirical analysis. *Journal of Statistical Mechanics: Theory and Experiment*, 2017(2):023406, 2017.

B. S. Bernanke, J. Boivin, and P. Eliasz. Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach. *The Quarterly Journal of Economics*, 120(1):387–422, 2005.

F. Black. Noise. *The Journal of Finance*, 41(3):528–543, 1986.

S. Bonhomme and J.-M. Robin. Consistent noisy independent component analysis. *Journal of Econometrics*, 149(1):12–25, 2009.

J.-P. Bouchaud, Y. Gefen, M. Potters, and M. Wyart. Fluctuations and response in financial markets: the subtle nature of 'random' price changes. *Quantitative Finance*, 4(2):176–190, 2004.

J.-P. Bouchaud, J. D. Farmer, and F. Lillo. How markets slowly digest changes in supply and demand. In *Handbook of Financial Markets: Dynamics and Evolution*, pages 57–160. Elsevier, 2009.

M. T. Bradshaw. How do analysts use their earnings forecasts in generating stock recommendations? *The Accounting Review*, 79(1):25–50, 2004.

J. Brogaard, A. Carrion, T. Moyaert, R. Riordan, A. Shkilko, and K. Sokolov. High frequency trading and extreme price movements. *Journal of Financial Economics*, 128 (2):253 – 265, 2018.

L. D. Brown, A. C. Call, M. B. Clement, and N. Y. Sharp. Inside the Black Box of sell-side financial analysts. *Journal of Accounting Research*, 53(1):1–47, 2015.

M. Brunnermeier and L. Pedersen. Predatory trading. *The Journal of Finance*, 60(4): 1825–1863, 2005.

F. Bucci, I. Mastromatteo, Z. Eisler, F. Lillo, J.-P. Bouchaud, and C.-A. Lehalle. Co-impact: crowding effects in institutional trading activity. *Quantitative Finance*, 20(2): 193–205, 2020.

P. W. Buchen and M. Kelly. The maximum entropy distribution of an asset inferred from option prices. *Journal of Financial and Quantitative Analysis*, 31(1):143–159, 1996.

L. M. Calcagnile, G. Bormetti, M. Treccani, S. Marmi, and F. Lillo. Collective synchronization and high frequency systemic instabilities in financial markets. *Quantitative Finance*, 18(2):237–247, 2018.

J.-F. Cardoso. High-order contrasts for independent component analysis. *Neural Computation*, 11(1):157–192, 1999.

M. M. Carhart. On persistence in mutual fund performance. *The Journal of Finance*, 52 (1):57–82, 1997.

B. Carlin, M. Lobo, and S. Viswanathan. Episodic liquidity crises: Cooperative and predatory trading. *The Journal of Finance*, 65(5):2235–2274, 2007.

C. Casey. Corporate valuation, capital structure and risk management: A stochastic DCF approach. *European Journal of Operational Research*, 135(2):311–325, 2001.

CFTC-SEC. *Findings regarding the market events of May 6, 2010.* Report, 2010.

E. C. Chang, Y. Luo, and J. Ren. Pricing deviation, misvaluation comovement, and macroeconomic conditions. *Journal of Banking & Finance*, 37(12):5285–5299, 2013.

T. Chordia, R. Roll, and A. Subrahmanyam. Commonality in liquidity. *The Journal of Financial Economics*, 56(1):3–28, 2000.

A. Coad and R. Rao. Firm growth and R&D expenditure. *Economics of Innovation and New Technology*, 19(2):127–145, 2010.

P. Comon. Independent component analysis, a new concept? *Signal Processing*, 36(3): 287–314, 1994.

H. Cramér. *Random variables and probability distributions*, volume 36. Cambridge University Press, 2004.

D. Cutler, J. Poterba, and L. Summers. What moves stock prices? *The Journal of Portfolio Management*, 15(3):4–12, 1989.

A. Damodaran. Valuation approaches and metrics: A survey of the theory and evidence. *Foundations and Trends® in Finance*, 1(8):693–784, 2007.

A. Damodaran. *Investment valuation: Tools and techniques for determining the value of any asset.* John Wiley & Sons, 2012.

G. Darmois. Analyse générale des liaisons stochastiques: etude particulière de l'analyse factorielle linéaire. *Revue de l'Institut International de Statistique*, pages 2–8, 1953.

D. Dayananda, R. Irons, S. Harrison, J. Herbohn, and P. Rowland. *Capital budgeting: financial appraisal of investment projects*. Cambridge University Press, 2002.

C. Doz, D. Giannone, and L. Reichlin. A quasi-maximum likelihood approach for large, approximate dynamic factor models. *The Review of Economics and Statistics*, 94(4): 1014–1024, 2012.

J. Durbin and S. J. Koopman. *Time series analysis by state space methods*, volume 38. Oxford University Press, 2012.

J. Eriksson and V. Koivunen. Identifiability, separability, and uniqueness of linear ICA models. *IEEE Signal Processing Letters*, 11(7):601–604, 2004.

R. C. Fair. Events that shook the market. *The Journal of Business*, 75(4):713–731, 2002.

E. F. Fama and K. R. French. The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465, 1992.

E. F. Fama and K. R. French. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56, 1993.

E. F. Fama and K. R. French. A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22, 2015.

E. F. Fama and J. D. MacBeth. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3):607–636, 1973.

M. Forni, M. Hallin, M. Lippi, and L. Reichlin. The generalized dynamic-factor model: Identification and estimation. *The Review of Economics and Statistics*, 82(4):540–554, 2000.

M. Forni, D. Giannone, M. Lippi, and L. Reichlin. Opening the black box: Structural factor models with large cross sections. *Econometric Theory*, 25(5):1319–1347, 2009.

M. Forni, L. Gambetti, M. Lippi, and L. Sala. Noisy news in business cycles. *American Economic Journal: Macroeconomics*, 9(4):122–52, 2017.

N. French and L. Gabrielli. Discounted cash flow: accounting for uncertainty. *Journal of Property Investment & Finance*, 23(1):75–89, 2005.

P. Gagliardini and C. Gouriéroux. Identification by Laplace transforms in nonlinear time series and panel models with unobserved stochastic dynamic effects. *Journal of Econometrics*, 208(2):613–637, 2019.

L. C. Garcia del Molino, I. Mastromatteo, M. Benzaquen, and J.-P. Bouchaud. The multivariate Kyle model: More is different. *SIAM Journal on Financial Mathematics*, 11(2):327–357, 2020.

N. Gârleanu and L. H. Pedersen. Dynamic trading with predictable returns and transaction costs. *The Journal of Finance*, 68(6):2309–2340, 2013.

D. Gimpelevich. Simulation-based excess return model for real estate development. *Journal of Property Investment & Finance*, 2011.

A. Golub, J. Keane, and S.-H. Poon. High frequency trading and mini flash crashes. *Available at SSRN 2182097*, 2012.

G. H. Golub and C. F. Van Loan. *Matrix Computations*. JHU Press, fourth edition, 2013.

M. J. Gordon. The savings investment and valuation of a corporation. *The Review of Economics and Statistics*, 44(1):37–51, 1962.

C. Gouriéroux, A. Monfort, and J.-P. Renne. Statistical inference for independent component analysis: Application to structural VAR models. *Journal of Econometrics*, 196 (1):111–126, 2017.

C. Gouriéroux, A. Monfort, and J.-P. Renne. Identification and estimation in non-fundamental structural VARMA models. *The Review of Economic Studies*, 87(4): 1915–1953, 2019.

B. Graham and D. L. Dodd. *Security Analysis*. Whittlesey House, McGraw-Hill Book Company, 1934.

E. Granziera, H. R. Moon, and F. Schorfheide. Inference for VARs identified with sign restrictions. *Quantitative Economics*, 9(3):1087–1121, 2018.

O. Guéant. *The Financial Mathematics of Market Liquidity: From optimal execution to market making*, volume 33. CRC Press, 2016.

M. Hallin and R. Liška. Determining the number of factors in the general dynamic factor model. *Journal of the American Statistical Association*, 102(478):603–617, 2007.

A. C. Harvey. *Forecasting, structural time series models and the Kalman filter*. Cambridge university press, 1990.

D. Hirshleifer and D. Jiang. A financing-based misvaluation factor and the cross-section of expected returns. *The Review of Financial Studies*, 23(9):3401–3436, 2010.

R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1990.

A. Hyvärinen. Fast independent component analysis for noisy data using gaussian moments. In *1999 IEEE International Symposium on Circuits and Systems (ISCAS)*, 1999.

A. Hyvärinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. Wiley, 2001.

S. Ikeda. ICA on noisy data: A factor analysis approach. In *Advances in Independent Component Analysis*, pages 201–215. Springer, 2000.

N. Johnson, G. Zhao, E. Hunsader, H. Qi, N. Johnson, J. Meng, and B. Tivnan. Abrupt rise of new machine ecology beyond human response time. *Scientific Reports*, 3:2627, 2013.

M. Joho, H. Mathis, and R. Lambert. Overdetermined blind source separation: Using more sensors than source signals in a noisy mixture. In *Independent Component Analysis and Blind Signal Separation ICA 2000*, pages 81–86, 2000.

A. Joulin, A. Lefevre, D. Grunberg, and J.-P. Bouchaud. Stock price jumps: news and volume play a minor role. *arXiv preprint arXiv:0803.1769*, 2008.

B. Jungbacker and S. J. Koopman. Likelihood-based analysis for dynamic factor models. Technical report, Tinbergen Institute Discussion Paper, 2008.

A. Kirilenko, A. S. Kyle, M. Samadi, and T. Tuzun. The flash crash: The impact of high frequency trading on an electronic market. *The Journal of Finance*, 72(3):967–998, 2017.

T. Koller, M. Goedhart, and D. Wessels. *Valuation: measuring and managing the value of companies*, volume 499. John Wiley & Sons, 2010.

L. Kruschwitz and A. Löffler. *Discounted cash flow: a theory of the valuation of firms*. John Wiley & Sons, 2006.

A. S. Kyle. Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1335, 1985.

A. Lachapelle, J.-M. Lasry, C.-A. Lehalle, and P.-L. Lions. Efficiency of the price formation process in presence of high frequency participants: a mean field game analysis. *Mathematics and Financial Economics*, 10(3):223–262, 2016.

N. S. Lambert, M. Ostrovsky, and M. Panov. Strategic trading in informationally complex environments. *Econometrica*, 86(4):1119–1157, 2018.

M. Lanne, M. Meitz, and P. Saikkonen. Identification and estimation of non-gaussian structural vector autoregressions. *Journal of Econometrics*, 196(2):288–304, 2017.

O. Ledoit and M. Wolf. Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, 15(5):850–859, 2008.

F. Li, T.-M. Chow, A. Pickard, and Y. Garg. Transaction costs of factor-investing strategies. *Financial Analysts Journal*, 75(2):62–78, 2019.

X. Luo and A. Schied. Nash equilibrium for risk-averse investors in a market impact game with transient price impact. *Market Microstructure and Liquidity*, to appear, 2020.

H. Lütkepohl and A. Netšunajev. Structural vector autoregressions with heteroskedasticity: A review of different volatility models. *Econometrics and Statistics*, 1:2–18, 2017.

I. Mastromatteo, M. Benzaquen, Z. Eisler, and J.-P. Bouchaud. Trading lightly: Cross-impact and optimal portfolio execution. *Risk*, 30:82–87, 2017.

C. C. Moallemi, B. Park, and B. Van Roy. Strategic execution in the presence of an uninformed arbitrageur. *Journal of Financial Markets*, 15(4):361–391, 2012.

A. Moneta, D. Entner, P. O. Hoyer, and A. Coad. Causal inference by independent component analysis: Theory and applications. *Oxford Bulletin of Economics and Statistics*, 75(5):705–730, 2013.

E. Moulines, J.-F. Cardoso, and E. Gassiat. Maximum likelihood for blind separation and deconvolution of noisy signals using mixture models. In *Acoustics, Speech, and Signal Processing*, volume 5, pages 3617–3620, 1997.

A. A. Obizhaeva and J. Wang. Optimal trading strategy and supply/demand dynamics. *Journal of Financial Markets*, 16(1):1–32, 2013.

M. O'Hara. *Market Microstructure Theory*. Wiley, 1997.

A. Onatski. Determining the number of factors from empirical distribution of eigenvalues. *The Review of Economics and Statistics*, 92(4):1004–1016, 2010.

R. Razgaitis. *Valuation and dealmaking of technology-based intellectual property: Principles, methods and tools*. John Wiley & Sons, 2009.

R. Rigobon. Identification through heteroskedasticity. *The Review of Economics and Statistics*, 85(4):777–792, 2003.

M. Samis and G. A. Davis. Using Monte Carlo simulation with DCF and real options risk pricing techniques to analyse a mine financing proposal. *International Journal of Financial Engineering and Risk Management*, 1(3):264–281, 2014.

A. Schied and T. Zhang. A state-constrained differential game arising in optimal portfolio liquidation. *Mathematical Finance*, 27(3):779–802, 2017.

A. Schied and T. Zhang. A market impact game under transient price impact. *Mathematics of Operations Research*, 44(1):102–121, 2018.

A. Schied, T. Schöneborn, and M. Tehranchi. Optimal basket liquidation for CARA investors is deterministic. *Applied Mathematical Finance*, 17(6):471–489, 2010.

M. Schneider and F. Lillo. Cross-impact and no-dynamic-arbitrage. *Quantitative Finance*, 19(1):137–154, 2019.

T. Schöneborn. *Trade execution in illiquid markets: Optimal stochastic control and multi-agent equilibria*. Doctoral thesis, Technische Universität Berlin, Fakultät II - Mathematik und Naturwissenschaften, Berlin, 2008.

E. Sentana and G. Fiorentini. Identification, estimation and testing of conditionally heteroskedastic factor models. *Journal of Econometrics*, 102(2):143–164, 2001.

W. F. Sharpe. The Sharpe ratio. *Journal of Portfolio Management*, 21(1):49–58, 1994.

R. J. Shiller. Do stock prices move too much to be justified by subsequent changes in dividends? *The American Economic Review*, 71(3):421–436, 1981.

S. Shimizu, P. O. Hoyer, A. Hyvärinen, and A. Kerminen. A linear non-gaussian acyclic model for causal discovery. *Journal of Machine Learning Research*, 7(10):2003–2030, 2006.

C. A. Sims. Macroeconomics and reality. *Econometrica*, 48(1):1–48, 1980.

V. P. Skitovitch. On a property of the normal distribution. *DAN SSSR*, 89:217–219, 1953.

F. A. Sortino and L. N. Price. Performance measurement in a downside risk framework. *The Journal of Investing*, 3(3):59–64, 1994.

F. Steiger. The validity of company valuation using discounted cash flow methods. *arXiv, 1003.4881*, 2010.

J. H. Stock and M. W. Watson. Understanding changes in international business cycle dynamics. *Journal of the European Economic Association*, 3(5):968–1006, 2005.

J. H. Stock and M. W. Watson. Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. In *Handbook of Macroeconomics*, volume 2, pages 415–525. Elsevier, 2016.

E. Strehle. Optimal execution in a multiplayer model of transient price impact. *Market Microstructure and Liquidity*, 3(4):1850007, 2017a.

E. Strehle. *Single- and multiplayer trade execution strategies under transient price impact*. Doctoral thesis, School of Business Informatics and Mathematics, Universität Mannheim, Mannheim, 2017b.

M. Tomas, I. Mastromatteo, and M. Benzaquen. Cross impact in derivative markets. *Available at SSRN 3779461*, 2021.

G. Tsoukalas, J. Wang, and K. Giesecke. Dynamic portfolio execution. *Management Science*, 65(5):2015–2040, 2019.

J. Viebig, T. Poddig, and A. Varmaz. *Equity valuation: models from leading investment banks*, volume 434. John Wiley & Sons, 2008.

M. Vodret, I. Mastromatteo, B. Tóth, and M. Benzaquen. A stationary Kyle setup: Microfounding propagator models. *arXiv preprint arXiv:2011.10242*, 2020.