

Ph.D Thesis

Options trading strategies and equity risk premia

Author:

Davide Tedeschini

Thesis Committee:

Prof. Paul Schneider - Swiss Finance Institute, Università della Svizzera italiana
Prof. Patrick Gagliardini - Swiss Finance Institute, Università della Svizzera italiana
Prof. Olivier Scaillet- Swiss Finance Institute, Université de Genève

A thesis submitted in fulfillment of the requirements for the degree of Ph.D in Finance at the
Swiss Finance Institute – Università della Svizzera italiana - Faculty of Economics

July, 2018

Contents

Introduction	3
Approximate arbitrage with limit orders	6
Trading central moments	44
Evaluating models jointly with economic and statistical criteria . .	75

Introduction

This doctoral thesis examines, from both a theoretical and an empirical perspective, different aspects of the equity derivative markets, such as the appropriate evaluation of equity risk premia and the development of trading strategies based on options. The thesis is predominantly proposing new estimation techniques that do not necessarily require high-frequency data to be fully implemented. Only when I explore the microstructure implications of liquidity shortage for the design of trading strategies, I test the theoretical findings using intra-day transaction prices.

I have started my doctoral studies with a deep investigation of equity risk premia, since a broad and growing literature in the field has documented that stocks can be exposed to multiple risk factors and carry multiple risk premia (Bakshi et al., 2003; Bondarenko, 2003; Bollerslev et al., 2009; Carr and Wu, 2009; Neuberger, 2012). The interconnection among the different risk compensations are scrutinized by my advisor Paul Schneider through the construction of the so called *likelihood ratio swap* (Schneider (2015)). Similar to variance swaps, that trade implied variance for realized variance, the likelihood ratio swap trades implied pricing kernel variance for realized pricing kernel variance. In a follow up paper entitled “Evaluating models jointly with economical and statistical criteria”, I examine the features of this instrument to directly trade the pricing kernel using a model and a panel of options. The resulting trading strategy prescribes portfolio weights depending explicitly on the model parameters and on the option-implied forward-neutral density, thereby relating the predictive density of a model to how much money could be made or lost with it in the market. It thus combines statistical and economic information, addressing the concerns raised in Leitch and Tanner (1991) about the possible divergence of statistical and economic predictability criteria.

Risk premia are interpreted as expected profits from trading strategies and computed ex post as the difference of two components: the price of a traded asset and its corresponding realized payoff. For assessing the first component, option-based measures have become extremely popular in the last decade and the information arising from this type of derivatives has been increasingly exploited for retrieving model-free quantities. The most well-known example in this context is given by the VIX index computed by the Chicago Board Options Exchange (CBOE). In 2003, the original methodology where the index

was calculated as a Black-Scholes implied volatility, was replaced, at first, by a model-free measure introduced by [Britten-Jones and Neuberger \(2000\)](#) and, subsequently, with the formula developed by [Carr and Madan \(2001\)](#), the most popular nowadays. Their basic approach permits to compute implied noncentral moments for a generic function of the asset returns. One of the widest used application of the latter work is the construction of an option-based measure for central moments, as shown by [Bakshi et al. \(2003\)](#). Concerning the second component of risk premia, several studies have identified realized measures for non central moments ([Andersen et al., 2001](#); [Barndorff-Nielsen, 2002](#)) but no equivalent measure was analyzed in the literature, to the best of my knowledge, to compute realized central moments. A second joint work with Paul Schneider, denominated “Trading Central Moments”, fills this gap by introducing realized moments which reflect perfectly the implied measures of [Bakshi et al. \(2003\)](#).

The two above-mentioned working papers rely primarily on options for the construction of implied measures and of trading strategies. Working on these two projects I have developed a strong interest in the functioning of derivative markets, especially at a more microstructural level, that has led me to start a new paper on the potential opportunities arising when liquidity shocks originate in order-driven markets. I have designed a trading strategy based on the placement of aggressive limit orders for otherwise identical options with different strikes or styles. I have analysed the performance of the strategy both theoretically and empirically and concluded that the scheme can be deemed as a practical example of *approximate arbitrage* ([Bernardo and Ledoit \(2000\)](#)), an investment opportunity with an exceptionally high ratio between the expected positive and the expected negative payoff. These findings are illustrated in the paper entitled “Approximate arbitrage with limit orders”.

References

- Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens (2001). The distribution of realized stock return volatility. *Journal of financial economics* 61(1), 43–76.
- Bakshi, G., N. Kapadia, and D. Madan (2003). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies* 16(1), 101–143.
- Barndorff-Nielsen, O. E. (2002). Econometric analysis of realized volatility and its use in

- estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64(2), 253–280.
- Bernardo, A. E. and O. Ledoit (2000). Gain, loss, and asset pricing. *Journal of political economy* 108(1), 144–172.
- Bollerslev, T., G. Tauchen, and H. Zhou (2009). Expected stock returns and variance risk premia. *Review of Financial Studies* 22(11), 4463–4492.
- Bondarenko, O. (2003). Why are put options so expensive. *University of Illinois at*.
- Britten-Jones, M. and A. Neuberger (2000). Option prices, implied price processes, and stochastic volatility. *The Journal of Finance* 55(2), 839–866.
- Carr, P. and D. Madan (2001). Towards a theory of volatility trading. *Option Pricing, Interest Rates and Risk Management, Handbooks in Mathematical Finance*, 458–476.
- Carr, P. and L. Wu (2009). Variance risk premiums. *Review of Financial Studies* 22(3), 1311–1341.
- Leitch, G. and J. E. Tanner (1991). Economic forecast evaluation: profits versus the conventional error measures. *The American Economic Review*, 580–590.
- Neuberger, A. (2012). Realized skewness. *Review of Financial Studies* 25(11), 3423–3455.
- Schneider, P. (2015). Generalized risk premia. *Journal of Financial Economics* 116(3), 487–504.

Approximate arbitrage with limit orders*

Davide Tedeschini

Swiss Finance Institute - University of Lugano

Abstract

Almost riskless investment opportunities represent a fundamental innovation of the recent developments in asset pricing theory. In this paper, I introduce a related trading scheme involving two options and two asynchronous operations: a limit order for one of the assets and a market order for the other one, once the limit order is executed. A model integrating option pricing and order arrivals explains the proximity of this strategy to a pure arbitrage. In particular, satisfying the requisites of the *approximate arbitrage opportunities*, I therefore refer to it as a *limit order approximate arbitrage*. An empirical study on a novel option data set confirms that market participants actively invest in these trades. The analysis also reveals the presence of short-living pure arbitrage opportunities in the market, promptly taken by the arbitrageurs.

JEL Classification: G12; G13; G17.

Keywords: Trading strategies, order book, options, Markov chains.

*I have benefited from helpful discussions on this topic with Harjoat Bhamra, Christian Dorion, Steve Heston, Steven Kou, Lorian Mancini, Dmitriy Muravyev, Mirela Sandulescu, Paul Schneider and the seminar participants at University of Maryland and at University of Lugano. Financial support from Swiss National Science Foundation (project “Equity risk premia: model evaluation, trading strategies and estimation of implied moments”) is gratefully acknowledged. Part of this paper was written while I was visiting Robert Smith Business School at University of Maryland. Contact at davide.tedeschini@usi.ch, University of Lugano, Via Buffi 13, 6900 Lugano.

1 Introduction

Asset pricing theory has introduced different categories of almost riskless investment opportunities, such as approximate and statistical arbitrages (see e.g. [Bernardo and Ledoit \(2000\)](#) and [Bondarenko \(2003\)](#)). These strategies differ from pure arbitrages, intended as fully riskless investments, because of the small, but still positive, probability of negative payoffs. The theoretical implications of their existence in complete and incomplete markets has been extensively investigated in a successful attempt to restrict the set of admissible stochastic discount factors. Nonetheless, little is known about how to design these types of schemes in practice. As the activity of the arbitrageurs is pivotal for removing law-of-one price violations, the presence of traders exploiting almost riskless strategies is essential for the validity of all the related pricing restrictions. Several concerns then naturally arise when examining tangible instances of approximate and statistical arbitrages: how often do these opportunities originate in financial markets? If they appear, what economic conditions motivate traders to immediately grab them? Which implementation issues might they tackle? This paper addresses these questions by constructing an approximate arbitrage on two otherwise identical options differing either by strikes or by styles.

The investor employs aggressive limit orders, i.e. orders near the Bid-Ask spread, to develop the approximate arbitrage scheme. Hence, I denote these strategies as *limit order approximate arbitrages* (LOAA henceforth), where an *approximate arbitrage opportunity* is a zero cost investment strategy offering an exceptionally high “gain-loss” ratio, computed dividing the expected positive payoff by the expected negative payoff. Indeed, no initial expenditure is required to enter in the strategy and out of four possible final states, only one entails a loss. However, corresponding to twice the fee paid to submit the limit order, the amount of the loss is relatively modest. Consequently, the gain-loss ratio is extremely high, so that the strategy qualifies as an approximate arbitrage.

In a hypothetical ranking among types of investment opportunities based on the gain-loss ratio, approximate arbitrages would be the closest ones to fully riskless strategies, whose ratio is equal to infinite. Then, it is crucial to clearly distinguish the two sets of opportunities in the context of this article. The functioning of a pure arbitrage

for two mispriced generic assets is rather straightforward. Let the two assets be H and L , such that no arbitrage forces the Ask price of H to be higher or equal than the Bid price of L . When arbitrageurs detect a violation of the law-of-one price between the two assets ($L^{Bid} > H^{Ask}$), they place two market orders to earn an immediate riskless profit. A broad literature in this field therefore focuses on trades occurring at opposite sides of the order book. This paper instead sheds light on the case $L^{Ask} > H^{Ask}$. Being the Bid-Ask spread normally positive, this type of market condition is intrinsically more likely to occur than a pure arbitrage violation.

I propose a simple model to analyse this trading scheme and its profitability under three perspectives. First, I compare its gain-loss ratio with the same measure computed for both an out of the money option and for the underlying asset, to certify that the strategy is indeed an approximate arbitrage. For reasonable model specifications, the level of the ratio oscillates between three and ten, meaning that the area under the positive part of the payoff distribution is from three to ten times bigger than its negative counterpart. Second, I demonstrate that a logarithmic utility investor finds implementing the strategy worthwhile in almost all market conditions. In fact, she does not submit the aggressive limit order only for overly unfavourable model parametrizations. Third, by simulating the Bid and Ask prices of the two options, I also gauge the strategy returns and highlight that their level is strikingly high. This result is not surprising, knowing that approximate arbitrages are investment opportunities lying in the *neighbourhood* of pure arbitrages.

An empirical study of the Eurex derivative market supports the idea of traders actually investing in LOAAs. Analysing nano-second synchronized option transactions, I also identify several pure arbitrage trades for options with different strikes and styles. The latter findings are particularly noticeable, since trading venues where both American and European-style options are exchanged represent an uncommon peculiarity. Along with the London Stock Exchange (LSE), the Eurex is the only market, to the best of my knowledge, where otherwise identical American and European-style options are simultaneously traded. More specifically, both the types of derivatives can be written on the FTSE100 index in the LSE, and on twenty-eight individual stocks in the Eurex market.

After unveiling the existence of arbitrage trades in the option market, I develop a

simple method for a better understanding of the agents behind these trades. For an arbitrage trade to be profitable, the gain should cover the transaction costs needed to execute it. Hence, I analyse the order fees in the Eurex market for the three types of participants who are allowed to open a trading account: brokers, dealers and market makers. Results suggest that brokers and dealers would have not had any incentive in implementing 91% and 70% of the arbitrages respectively, indicating that most of the opportunities could have been profitable only for market makers. Thus, their ability to establish a transaction cost advantage over other investors permit them to exploit these free lunches.

Related Literature. The term *approximate arbitrage* dates back to [Shanken \(1992\)](#), who defines it as an asset having more than twice the market Sharpe ratio. This criterion is also adopted by [Cochrane and Saa-Requejo \(2000\)](#) in the formalization of the so called “good deals”, i.e. investment opportunities with extraordinarily high Sharpe ratios. However, [Bernardo and Ledoit \(2000\)](#) document how to devise arbitrage opportunities with low Sharpe Ratio, recommending the usage of the gain-loss ratio to determine a set of strategies adjacent to pure arbitrages. Statistical arbitrages can also be included in this set. Theorized by [Bondarenko \(2003\)](#), they might result in negative outcomes even if the average payoff in each final state should be non-negative. I show that LOOAs do not satisfy this property, because the trajectory leading to the state implying a loss is unique.

The profitability of LOOAs is investigated in a model combining full-fledged option pricing frameworks, such as [Black and Scholes \(1973\)](#) and [Heston \(1993\)](#), with the insights of a new branch of market microstructure studying order-driven markets in a dynamic setting. In particular, the limit order book is modelled in an environment with no asymmetric information. [Parlour \(1998\)](#) provides a first instance of the agent’s decision process in submitting limit or market orders. [Roşu \(2009\)](#) extends the analysis by allowing for cancellation orders and [Cont et al. \(2010\)](#) assume the arrival of orders in the book to be completely exogenous. I borrow this assumption and formalize a state space representation of the Bid-Ask spreads, similar to the one of [Foucault et al. \(2005\)](#), to describe an order-driven market where two options are traded.

By identifying distinct examples of pure arbitrage trades, this article also contributes to the literature on arbitrage violations. In this regard, [Grossman and Stiglitz \(1976\)](#),

1980) were the first ones to point out that, if all traders presume that arbitrage opportunities will never be spotted, the motivation to investigate financial transactions will be wiped out. Lack of focus on trades may then originate arbitrage windows, leading to the so called “arbitrage paradox”. As already argued by Akram et al. (2008) in the context of foreign exchange markets, until high-frequency data became accessible to both investors and researchers, the needle was hard to be threaded. Indeed, the active role of arbitrageurs in removing mispricing is difficult to detect with daily or monthly data, because of the short-lived nature of arbitrages.

The existence of arbitrage violations in financial markets has been examined for a broad set of asset classes, including equity, futures, options, currencies and credit spreads. Concerning options, put-call parity is one of the most tested no arbitrage constraint. Stoll (1969) and Gould and Galai (1974) were among the first studies to document overpricing and underpricing for both puts and calls. Yet, their analysis involves American options, when the put-call parity holds as an equality only for European options. This issue is addressed by Kamara and Miller (1995), who use intraday data to show that, once the effect of early exercise premium is taken away, the frequency and size of put-call parity violations drop substantially.

Short-living riskless profits are reported also in the foreign exchange markets, both as violations of the covered interest rate parity relation (Rhee, Rhee; Akram et al., 2008) and as triangular arbitrages (Aiba and Hatano, 2004; Foucault et al., 2017). Index arbitrage represents another example to be included in this branch of research. MacKinlay and Ramaswamy (1988) and Stoll and Whaley (1990) examine the profitability of index arbitrage. The main weakness in their approach consists in selecting the reported index as a proxy of the true one. Chung (1991) considers the true value of the index to empirically demonstrate that the magnitude and frequency of profitable opportunities are appreciably smaller.

The paper is organized according to the following outline. In Section 2, after illustrating a concrete instance of LOAA, I introduce the model. Section 3 focuses on the simulation part, Section 4 is devoted to the empirical study of the Eurex option market and Section 5 concludes. Appendix A contains an additional example on the approximate arbitrage implementation. Proofs, figures and a detailed description of the novel dataset are included in Appendix B, Appendix C and Appendix D respectively.

2 The strategy in the context of option markets

2.1 Definition and example

I introduce the notion of limit order approximate arbitrages in a market with two otherwise identical call options with different strikes or styles. To frame these two characteristics in a general setting, I define $C_H^{Bid}(t, T)$, $C_L^{Bid}(t, T)$, $C_H^{Ask}(t, T)$ and $C_L^{Ask}(t, T)$ as the Bid and Ask prices of a *high* (H) and *low* (L) call options. Then, I distinguish the two features as follows:

1. **Strikes:** for two generic strikes $K1 > K2$, no arbitrage implies $C_H^{Ask}(t, T) = C_{K2}^{Ask}(t, T) > C_{K1}^{Bid}(t, T) = C_L^{Bid}(t, T)$. The opposite relation holds for put options.
2. **Styles:** for European (E) and American-style (A) options, no arbitrage implies $C_H^{Ask}(t, T) = C_A^{Ask}(t, T) > C_E^{Bid}(t, T) = C_L^{Bid}(t, T)$. The same inequality holds for put options as well.

A precise characterization of a LOAA permits to highlight, subsequently, the main differences with a pure arbitrage. I qualify them as “approximate” in the sense of [Bernardo and Ledoit \(2000\)](#) and justify this choice in Section 2.2. According to their definition, an approximate arbitrage provides an extraordinary high ratio between potential profit and loss of a financial strategy.

Definition 2.1. Limit order approximate arbitrage. A limit order approximate arbitrage is a zero-investment trading strategy where an investor:

1. Places a buy (sell) limit order for H (L) at a price $C_H^{Bid}(t, T) < C_L^{Bid}(t, T)$ ($C_L^{Ask}(t, T) > C_H^{Ask}(t, T)$).
2. If the limit order is executed, sells (buys) contemporaneously the option L (H) at a the price $C_L^{Bid}(t, T)$ ($C_H^{Ask}(t, T)$), earning the price difference $C_L^{Bid}(t, T) - C_H^{Bid}(t, T)$ ($C_L^{Ask}(t, T) - C_H^{Ask}(t, T)$).

The strategy is zero-investment since I suppose that the investor borrows, at an annual interest rate r , the amount of money corresponding to the order fee, that she will return at the expiration T .

A pure arbitrage opportunity (PAO), by contrast, is a zero-cost trading strategy that offers the possibility of a gain with no possibility of a loss. In the context of this study they materialize in option markets as fully riskless and self-financed opportunities.

Definition 2.2. Pure arbitrage. A *pure arbitrage trade* is the transaction executed by an arbitrageur when:

$$C_L^{Bid}(t, T) - C_H^{Ask}(t, T) - TC > 0, \quad (1)$$

where TC are the transaction costs needed to implement the arbitrage.

Transaction costs play a pivotal role in this field since theories on “limits of arbitrage” (Shleifer and Vishny (1997) and Gromb and Vayanos (2010) among others) postulate that, in presence of these opportunities, sophisticated investors such as hedge funds, may not be able to undo the mispricings because of the high transaction costs they are going to face. The execution costs of strike and style arbitrages consist only in the order book fees. Once the violation $C_H^{Ask}(t, T) < C_L^{Bid}(t, T)$ arises, an arbitrageur removes it by placing a sell market order for $C_L^{Bid}(t, T)$ and a buy market order for $C_H^{Ask}(t, T)$.

A toy example is described here to explain when LOAAs originate and how a trader can attain a profit from investing in them. Like in pure arbitrage trades, the investor constructs a position with final non-negative payoff by selling the option L at a higher price than the one needed to purchase the option H . However, LOAAs differ from pure arbitrages because they do not emerge as riskless opportunities, and because the two transactions occur at the same side of the order book (either Ask or Bid).

Let $\phi = 0.05$ be the percentage fee associated to every placeable order in the exchange, $r = 0.02$ be the annual interest rate, and $C_{100}^{Bid}(t, T)$, $C_{100}^{Ask}(t, T)$, $C_{110}^{Bid}(t, T)$ and $C_{110}^{Ask}(t, T)$ be the Bid and Ask prices of two European call options differing only in the level of the strikes, 100 and 110 respectively. They are traded in the same market and, at a random time t_M , the Bid-Ask spread of $C_{100}(t_M, T)$ is included in the one of $C_{110}(t_M, T)$. Table 1 shows an example of such a condition in a hypothetical order book. An arbitrageur does not observe any violation, being $C_{110}^{Bid}(t_M, T) = 0.7 < C_{100}^{Ask}(t_M, T) = 1.2$. Yet, she pursues the following strategy: at time t_{M+1} shortly after t_M , she places a sell order $C_{110}^{Ask}(t_{M+1}, T) = 1.3$ in-between the

$C_{100}(t_M, T)$		$C_{110}(t_M, T)$		$C_{100}(t_{M+1}, T)$		$C_{110}(t_{M+1}, T)$	
Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask
0.7		0.6		0.7		0.6	
0.8		0.7		0.8		0.7	
	1.2		1.4		1.2		1.3
	1.4				1.4		1.4

Table 1: Strategy at the Ask price. An example of the strategy when the Bid-Ask spread of $C_{100}(t_M, T)$ is included in the one of $C_{110}(t_M, T)$. The agent limit order is highlighted in bold.

two options best Asks, paying a fee of 0.065. In case a buy market order is posted afterwards, matching her limit order before a market/cancellation order for L arrives, the arbitrageur buys one option for 1.2, realizing an immediate profit of $1.3 - 1.2$ in t_{M+1} . Otherwise, the strategy is unsuccessful and she cancels the order.

The positive difference between the two Ask prices does not embody the sole gain of the strategy. The constructed position is in fact known as “bull call spread”, a type of vertical spread whose final payoff BCS_T resembles the one of a standalone call option with a ceiling on the profits:

$$BCS_T = \begin{cases} 0 & \text{if } S_T < 100, \\ S_T - 100 & \text{if } 100 < S_T < 110, \\ 10 & \text{if } S_T > 110, \end{cases} \quad (2)$$

where S_T is the value of the underlying at the expiration T . The potential losses related to the strategy are the two order fees (one limit order and one market order) and, possibly, the cancellation fee.

An analogous case can be conceived when the Bid-Ask spread of $C_{110}(t_M, T)$ is contained in the one of $C_{100}(t_M, T)$. This type of strategy is treated in Appendix A. Similar examples could be designed for the put options as well. Since the put option price is increasing in the strike, the only variation would reside in the position on the single options: short for the put with the higher strike and long for the one with the lower strike.

The model presented in the next section concentrates on strategies exploiting differences in the Ask price as in the example above. Indeed, given the direct empirical

relation between liquidity and moneyness documented by [Peña et al. \(2001\)](#), the spread of the call option with the higher moneyness (H) is commonly the one with the smaller spread. The same consideration holds for options with different styles, since as shown in [Section 4](#), when both styles are traded on the same underlying, American options are far more liquid than their European counterparts.

2.2 A model for option Bid and Ask prices

This section presents a model to accommodate the problem sketched in the previous example. To dissect LOAAs in a general framework, I introduce a market organized as a limit order book without intermediaries, where the two aforementioned European call options H and L are exchanged. To ease the notation, I define several quantities illustrated in the model considering a generic option $O = \{H, L\}$.

An investor I with initial wealth W and utility function $u(W)$ operates in the market. She formulates her views on the payoffs option probabilities adopting an option pricing model $\mathcal{M}(\theta)$, where θ is the vector of parameters. The probability at time t that H will be in-the-money in T is denoted by p_O^{ITM} . Consistently, $p_O^{OTM} = 1 - p_O^{ITM}$ is the probability that O will not be in-the-money. Keeping in mind that H is either the call with lower strike or the American option, if L will be in the money, so will H .

Market participants can place three types of orders: limit, market and cancellation orders. In the same spirit of [Foucault et al. \(2005\)](#) and [Cont et al. \(2010\)](#), all orders arrive at exponential times. In particular, I hypothesize that:

- Limit buy and sell orders enter in the book as best Ask (Bid) with rate λ_O .
- Market buy and sell orders arrive with rate μ_O .
- Cancellation orders for the best Bid and Ask arrive with rate γ_O .
- The above events are mutually independent.

Arrival intensities do not differ between buy and sell orders. The order fee is a percentage ϕ of the price, regardless of the type of order and of the participant issuing it. The standard time and price priority rules apply and all the limit orders are submitted for a unitary quantity, with δ being the minimum tick size.

Given these assumptions, the option Bid and Ask prices are composed of two variables: the fundamental value and the spread. The fundamental value $FV^O := C_O^{\mathcal{M}(\theta)}(t, T)$ is the price implied by the option pricing model, whereas the spreads SP^O is assumed to follow a Markov chain state space model with q states, where δq is the highest spread observable in the market and δ is the smallest one. Thus, the states of the chain are $i = 1, 2 \dots q$, and the possible values of the spreads are $j = \delta, 2\delta \dots q\delta$.

When the spread is equal to δ , limit orders inside it are excluded. No trader will in fact issue a limit buy (sell) order inside the spread for the same price of the Bid (Ask) limit order already present in the order book. A rational trader will instead post a buy (sell) market order, since she can purchase (sell) the option immediately for the desired price. Accordingly, since in this state only market and cancellation orders can be placed, the spread will certainly move from state 1 to state 2.

On the contrary, when the spread achieves its maximum level δq , I formalise a mechanism that guarantees an upper bound: if market and cancellation orders arrive, market makers immediately intervene by reinserting the matched or cancelled limit order that restore the spread at the level δq . Consequently, once in state q , the spread can either stay in the same state, if a market/cancellation order arrives, or move back to the state $q - 1$, if a limit order enters in the book. In all other states ($2 \dots q - 1$), the spread can either shift to a state $i + 1$, if a market/cancellation order is placed, or to a state $i - 1$, if a limit order is posted.

To construct the transition matrix of the Markov chain it is necessary to determine the probability p_L that a limit order arrives before both a market and a cancellation order.

Proposition 2.3. *The probability p_L is given by the following expression:*

$$p_L = \frac{\lambda_O}{\lambda_O + \mu_O + \gamma_O}. \quad (3)$$

Proof. In Appendix B. □

Being $1 - p_L$ the probability that a market/cancellation order will arrive before a

limit order, the transition matrix has then the following structure:

$$Y := \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ p_L & 0 & 1 - p_L & \dots & 0 & 0 \\ 0 & p_L & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & p_L & 1 - p_L \end{bmatrix},$$

where the k^{th} entry in the h^{th} row is the probability that the size of the spread moves to δk from δh , with $k, h = 1, 2, \dots, q$. Integrating the option pricing framework $\mathcal{M}(\theta)$ with the state space model of the spreads, the Ask and Bid prices are expressed by the equations:

$$\begin{aligned} C_H^{Ask}(t, T) &= C_H^{\mathcal{M}(\theta)}(t, T) + SP^H/2 & C_H^{Bid}(t, T) &= C_H^{\mathcal{M}(\theta)}(t, T) - SP^H/2 \\ C_L^{Ask}(t, T) &= C_L^{\mathcal{M}(\theta)}(t, T) + SP^L/2 & C_L^{Bid}(t, T) &= C_L^{\mathcal{M}(\theta)}(t, T) - SP^L/2. \end{aligned} \quad (4)$$

The agent I intends to construct the trading scheme outlined in Section 2.1. To achieve this goal, she monitors continuously the order book, waiting for a random time t_M when the Ask price of H will be lower than the Ask price of L :

$$t_M = \inf \{t \leq T : C_L^{Ask}(t, T) > C_H^{Ask}(t, T)\} \quad (5)$$

Depending on the width of this difference, it is possible to further recognize two sub-cases:

1. $C_L^{Ask}(t_M, T) - C_H^{Ask}(t_M, T) > \delta$. Then, the agent submits an aggressive sell limit order $C_L^{Ask, I}(t_M, T) = C_H^{Ask}(t_M, T) + \delta$. The choice is motivated by her incentive to increase the probability of the limit order to be matched by a market order. For example, in the case $C_L^{Ask}(t_M, T) - C_H^{Ask}(t_M, T) = 5\delta$, the investor could potentially place an order $C_L^{Ask, I}(t_M, T) = C_H^{Ask}(t_M, T) + 4\delta$ and implement the strategy with a higher immediate profit. Unfortunately, such an order is less likely to result in a trade than $C_L^{Ask, I}(t_M, T) = C_H^{Ask}(t_M, T) + \delta$. In summary, I conjecture that the arbitrageur risk aversion urges her to prefer a smaller, but more likely gain, to a higher but more uncertain one.

2. $C_L^{Ask}(t_M, T) - C_H^{Ask}(t_M, T) = \delta$. If the arbitrageur wishes to invest in the strategy, she is forced to issue a limit order $C_L^{Ask, I}(t_M, T) = C_H^{Ask}(t_M, T)$, realizing only a potential payoff at the expiration and no immediate gain.

The model aims at addressing the following question: when should the agent I submit the aggressive limit order? A plausible answer relies on an accurate evaluation of the initial costs and of the potential payoffs. The initial cost is merely given by the order fee $-\phi C_L^{Ask, I}(t_M, T)$ linked to the sell limit order, which is completely financed by borrowing the equivalent amount of money at an annual interest rate r . I denote as $V := \phi C_L^{Ask, I}(t_M, T)(1 + r(T - t_M))$ the value to pay back in T . The payoff structure is more elaborate. To fully assess it, I assume that two scenarios are possible, one successful, with probability p_S , and one unprofitable, with probability $1 - p_S$. The latter occurs when one of the following three orders arrives before the limit order $C_L^{Ask, I}(t_M, T)$ is executed:

- A market/cancel order that matches/removes the limit order $C_H^{Ask}(t_M, T)$. The potential profit vanishes and the investor cancels the order $C_L^{Ask, I}(t_M, T)$.
- A limit order issued by another trader for a price lower than $C_L^{Ask, I}(t_M, T)$. In this case, I assume that the agent perceives that the market is moving in the opposite direction and cancels the order $C_L^{Ask, I}(t_M, T)$.

In case of success, the agent buys the option H at the price $C_H^{Ask}(t_M, T)$ as soon as she executes the sell of C_L at price $C_L^{Ask, I}(t_M, T)$. The overall strategy payoff will be positive and will consist in the immediate gain and in vertical spread payoffs. In case of failure, the agents yields a negative payoff, equivalent to the cancellation fee $-\phi C_L^{Ask, I}(t_M, T)$. I define the immediate gain IMG as:

$$IMG := C_L^{Ask, I}(t_M, T) - C_H^{Ask}(t_M, T)(1 + \phi). \quad (6)$$

The fee is paid only for the buy market order of H , since the sell of L is executed at the arrival of a market order. As previously mentioned, this is the event that triggers the contemporaneous purchase of H . The realization of the options payoffs and their amount at time T strictly depend on the value of the underlying. Therefore, I denote

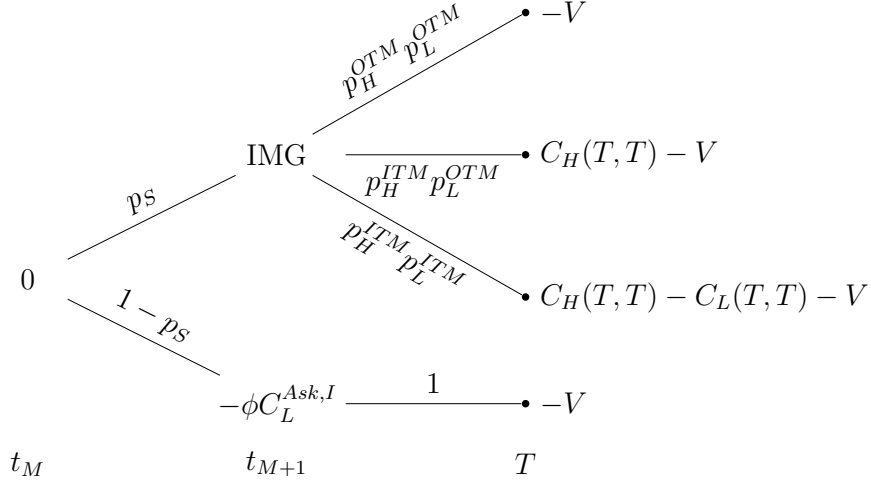


Figure 1: Limit order approximate arbitrage in a multinomial tree. The strategy is represented as a multinomial tree with three different times. At time t_M the agent places the sell order paying the fee $-\phi C_L^{Ask,I}(t_M, T)$. The evolution of the strategy depends first on the probability of success p_s and, eventually, on the vertical spread payoff at time T .

the strategy payoff as SP , a random variable structured as follows:

$$SP := \begin{cases} -\phi C_L^{Ask,I}(t_M, T) - V & \text{with } (1 - p_S), \\ IMG - V & \text{with } p_S p_H^{OTM} p_L^{OTM}, \\ IMG + C_H(T, T) - V & \text{with } p_S p_H^{ITM} p_L^{OTM}, \\ IMG + C_H(T, T) - C_L(T, T) - V & \text{with } p_S p_H^{ITM} p_L^{ITM}, \end{cases} \quad (7)$$

where $C_H(T, T)$ and $C_L(T, T)$ are the option payoffs.

The key reason behind the LOAA definition lies in the extremely small loss that the investor will face in case of failure. The net loss will be in fact equal to $-2\phi C_L^{Ask,I}(t_M, T)(1 + rT)$. Indeed, the agent pays the order fee twice, first to place the limit order in the book, and second to cancel the order. Figure 1 summarizes the evolution of the strategy in a multinomial tree. The scheme can be interpreted as an asset with no initial outlay and random payoffs realizable at two different subsequent times.

The agent submits the aggressive limit order if the expected utility of the trading scheme is higher than its current utility:

$$E_{t_M}[u(W + SP)] > u_{t_M}(W). \quad (8)$$

To calculate the expectation in Equation (8) the agent needs to estimate the three probabilities appearing in Equation (7). For the computation of p_H^{ITM} and p_L^{ITM} , I use the aforementioned model $\mathcal{M}(\theta)$. On the other hand, the probability of success p_S is given by:

$$p_S = \frac{\mu_L}{\mu_L + \lambda_L + \mu_H + \gamma_H}, \quad (9)$$

and the proof is tantamount to the one of Proposition 2.3.

Data on the order book are crucial for assessing the probability of success p_S . Exponential distribution parameters can be estimated as the number of orders divided by the total trading time in the sample (in minutes). Because every trade is a direct manifestation of a market order, their intensities μ_L and μ_H are the only ones obtainable also from data on trades.

3 Simulation

3.1 When investing in the scheme

The model illustrated before is applied to the practical example of Section 2.1 to document first, that LOOA achieve uncommonly high gain-loss ratio, and second, that a log-utility investor finds these strategies very appealing. To briefly sum up the example, the two call options, H and L , are traded with strike 100 and 110 respectively. Given the two Ask prices $C_L^{Ask}(t_M, T) = 1.4$, $C_H^{Ask}(t_M, T) = 1.2$, an aggressive sell limit order $C_L^{Ask, I}(t_M, T) = 1.3$, and an annual interest rate $r = 0.02$, the payoff SP is then:

$$SP := \begin{cases} -\phi 2.626 & \text{with } 1 - \frac{\mu_L}{\mu_L + \lambda_L + \mu_H + \gamma_H}, \\ 0.1 - \phi 2.513 & \text{with } \frac{\mu_L}{\mu_L + \lambda_L + \mu_H + \gamma_H} P_{\mathcal{M}(\theta)}(S_T < 100), \\ 0.1 - \phi 2.513 + S_T - 100 & \text{with } \frac{\mu_L}{\mu_L + \lambda_L + \mu_H + \gamma_H} P_{\mathcal{M}(\theta)}(100 < S_T < 110), \\ 0.1 - \phi 2.513 + 10 & \text{with } \frac{\mu_L}{\mu_L + \lambda_L + \mu_H + \gamma_H} P_{\mathcal{M}(\theta)}(S_T > 110), \end{cases} \quad (10)$$

where $P_{\mathcal{M}(\theta)}(\cdot)$ is the probability distribution implied by the option pricing model $\mathcal{M}(\theta)$. She formulates her views on the option payoff with both a Black-Scholes model and a Heston model. Overall, the framework contemplates parameters of diverse origin: the vector θ of the option pricing model, the four Poisson intensities μ_L , μ_H , λ_L and γ_H , and the commission percentage fee ϕ .

I compute the value of the gain-loss for the *LOAA*, for the underlying S_t and for the option L , assuming short-short selling of the latter two assets at time t , to render all the strategies zero investment opportunities. [Bernardo and Ledoit \(2000\)](#) define the gain-loss ratio *GLR* of a generic payoff x as:

$$GLR := \frac{E[\{x, 0\}^+]}{E[\{-x, 0\}^+]}. \quad (11)$$

To calculate it, I work with the payoff probabilities implied by a Black-Scholes model¹, with $\theta = \{\eta = 0.02, \sigma = 0.2\}$, and by a Heston model², with the quantities estimated by [Art-Sahalia et al. \(2007\)](#), namely $\theta = \{\eta = 0.02, \kappa = 5.07, \alpha = 0.0457, \epsilon = 0.048, \rho = -0.767\}$.

For assessing the exponential distribution intensities, I take advantage of the analysis of [Cont et al. \(2010\)](#). Working on the order book of stocks traded on the Tokyo Stock Exchange, they estimate a market order intensity equal to 0.94. They also retrieve different values for the limit and cancel order intensities, depending on the proximity of the order to the best Ask. Since I am interested in limit orders entering as best Ask, and in cancellation orders of the best Ask, I pick the intensities corresponding to the closest orders to this level. In summary, I select the following values: $\mu_L = \mu_H = 0.94$, $\lambda_L = 1.95$ and $\gamma_H = 0.71$. Eventually, the rate ϕ is chosen to be equal to 0.065, the midpoint between the reduced fee rate paid by a broker and the one paid by a dealer in the Eurex market (see Table 4).

Letting the parameters μ_L , μ_H and ϕ vary, Figure 2 and Figure 3 provide a graphical representation of the gain-loss ratios. The lowest level is the one of the underlying,

¹The underlying dynamics is given by $dS_t = \eta S_t dt + \sigma S_t dB_t$, where B_t is a Standard Brownian motion.

²The underlying dynamics is modelled as follows:

$$dS_t = \eta S_t dt + \sqrt{v_t} dB_t^1, \quad dv_t = \kappa(\alpha - v_t)dt + \epsilon dB_t^2, \quad (12)$$

where B_t^1 and B_t^2 are two Standard Brownian motions with $Cov(dB_t^1, dB_t^2) = \rho dt$.

whose value is fairly priced. Stochastic volatility impacts severely the *GLR* of the OTM option L , which, in this setting, has a larger chance to expire in the money than in the constant volatility environment of Black-Scholes. The LOAA features by far the highest ratio in the set of investment opportunities. Its value becomes extraordinarily high when ϕ is smaller than 0.05, since this parameter strongly affects the area on the negative part of the payoff distribution.

Having established that LOAA are outstandingly attractive, I investigate whether a logarithmic utility investor ³ I with an initial wealth $W = 100$ should submit the limit order $C_L^{Ask,I}(t_M, T) = 1.3$. The necessary precondition to pursue the strategy is given by:

$$\begin{aligned} \log(100) &< E_{t_M}(\log(100 + SP)) \\ &= \left(1 - \frac{\mu_L}{\mu_L + \lambda_L + \mu_H + \gamma_H}\right) \log(100 - \phi 2.626) + \frac{\mu_L}{\mu_L + \lambda_L + \mu_H + \gamma_H} \\ &\quad \times \left[\int_0^{100} \log(100.1 - \phi 2.513) dS_T + \int_{100}^{110} \log(0.1 - \phi 2.513 + S_T) dS_T \right. \\ &\quad \left. + \int_{110}^{\infty} \log(110.1 - \phi 2.513) dS_T \right], \end{aligned} \quad (13)$$

where the three integrals are computed with Montecarlo simulation.

The function $E_{t_M}(\log(100 + SP)) - \log(100)$ is displayed in Figures 4 and 5, where the positive region is evidently the one where the agent invests in the strategy. All the above-mentioned choices for the model parametrization are called “plausible parameters” and the curves in the left column of the two graphs are computed with this specification. In the set of plots exhibited in the right column, keeping fixed the values of the other parameters, I select $\lambda_L = 5$ and $\gamma_L = 5$ to represent an extreme unfavourable situation, since the probability of success p_S becomes remarkably small. However, even in such an exceptional case, only abnormal values of ϕ and μ_H discourage the agent from investing in the strategy, further justifying the denomination “limit order approximate arbitrages”.

³Results do not change substantially if assuming a risk-neutral investor.

3.2 Profits estimation

After having discussed the model parametrizations that lead the agent to invest in LOAAs, I assess their potential returns through a simulation of the Bid and Ask price trajectories for the two options. Starting at an initial time t , the prices of H and L are reproduced at every minute until the expiration. To this end, the value of the underlying asset is simulated in a Black-Scholes setting, together with the Markov chain evolution modelling the spreads.

Conditional on different values of the minimum thick δ and of the intensities λ_H and μ_L , I distinguish four different cases. All the other parameters are kept fixed to the one reported in the previous section. Bid-Ask spreads are denoted as homogeneous when the three order intensities are equivalent for the two options, and as non-homogeneous when this does not hold. As previously mentioned, [Peña et al. \(2001\)](#) find empirically a proportional relation between moneyness and Bid-Ask spreads, that I reproduce in the heterogeneous case; since market orders widen the spread and limit orders in between of Bid and Ask prices narrow it, I select a higher limit order rate arrival and a lower market order rate arrival for the option H ($\lambda_H > \lambda_L$ and $\mu_H < \mu_L$). Cancellation order intensities are chosen to be equal for both the options even in this case.

For every trajectory, the investor places a limit order at every time t_M such that $C_L^{Ask}(t_M, T) > C_H^{Ask}(t_M, T)$. Hence, in a single path, multiple manifestations of this event might occur. I compute the total number of these observations for every path, together with the number of times that entering in the strategy results in an immediate gain. To determine this value, I consider whether at time t_{M+1} both a market order for L and a limit order for H arrive. Indeed, such a case would imply that someone has purchased the option L at the price specified by the investor I , and that the option H has not been cancelled or matched by a market order. In other words, the strategy is successful and the agent constructs the vertical spread by grabbing an immediate gain. In addition, I identify the number of times an arbitrage violation originates ($C_L^{Bid}(t_m, T) > C_H^{Ask}(t_m, T)$) and the sum of the immediate gains, of the vertical spread payoffs and of the net profits. The average value for all these quantities are reported in [Table 2](#).

The number of investments in the strategy during the life of the two options is

exhibited in the first line. Not surprisingly, this value increases in the heterogeneous case, when a wider average spread for the option L increases the times when $C_L^{Ask}(t_m, T) > C_H^{Ask}(t_m, T)$. The same pattern is observed for the number of successes. Moreover, even if the model does not theoretically rule out arbitrage violations, no free lunch is detected in any of the simulated trajectories, because of the tiny value assigned to δ .

Average returns are remarkably high, showing once again the proximity of LOAAs to pure arbitrages. The return distribution are plotted in Figure 6 and is clearly bimodal. Being the two options out of the money at time t , for most of the paths the payoff of the vertical spread is null and only the immediate gain is collected. Its amount is relatively small and oscillates around the first peak of the distribution. The second peak is a direct consequence of paths in which at least the option L ends up in the money. These outcomes are indeed less likely, but could provide uncommonly high returns.

	Heterogeneous spreads		Homogeneous spreads	
	$\delta = 0.01$	$\delta = 0.005$	$\delta = 0.01$	$\delta = 0.005$
Strategies	6.34	5.35	4.98	3.99
Successes	0.76	0.64	0.45	0.34
Violations	0	0	0	0
Returns	76.42	149.22	251.22	240.19

Table 2: Simulated average values. The average number of strategies implemented, of successes, and of arbitrage violations is computed as percentage of the total observations in 6 months. Given $T = 6/12$, the prices $C_{100}^{Bid}(t, T)$, $C_{100}^{Ask}(t, T)$, $C_{110}^{Bid}(t, T)$ and $C_{110}^{Ask}(t, T)$ are simulated at every minute. Considering a day of seven hours and a month of 30 days, the total number of observation is equal to 75600. The underlying is simulated according to a GBM with $\mu = 0.02$, $\sigma = 0.2$ and starting value $S_t = 90$. The fundamental value is computed with the Black-Scholes formula. Fixing $\lambda_L = 1$ and $\mu_H = 1$, spreads are homogeneous when $\lambda_H = \mu_L = 1$ and heterogeneous when $\lambda_H = \mu_L = 0.8$. The other parameters are $\phi = 0.065$ and $\gamma_H = \gamma_L = 0.4$.

4 Empirics

4.1 The Eurex option market

The mechanism of LOAA trades demonstrates how the different liquidity of similar options can be exploited for generating almost-riskless profits. In this empirical inves-

tigation, I provide clean evidence of the implementation of this strategy in the Eurex exchange, the largest derivative market in Europe. It is a completely electronic limit order book market where options on indexes and stocks from more than ten European countries can be traded. Two types of orders are possible for both the option’s styles: limit orders and market orders. Limit orders can be “Good for a day” (GFD), “Good till canceled” (GTC) or “Good till date (GTD)”, if the order in the system is valid until the end of the day, until cancellation from the trader or until a specific date. The last possible limit order is the so called “Immediate or cancel” (GFD), if the market participant intends the order to be filled immediately. Market orders are not visible in the order book for any market participant and have no specific price limit, but are matched to the best available contra-side Bid or offer. Unless otherwise specified, every order is deemed as GFD, implying that most of the quotes are likely to be not older than one day. Hidden orders are not allowed, hence I can exclude the existence of transactions within the Bid-Ask spread.

The goal consists in identifying LOAAs and PAOs from option trades. For every transaction, Eurex specifies several attributes such as the Security ID, the traded price and size, the aggressor side, the option strike and expiration, and the official time of trade execution with nanosecond time precision. I distinguish two data samples, depending on whether the object is a strike or a style arbitrage.

The first sample, denoted as “strike”, concerns all the transactions from January 1st 2014 till May 23rd 2017 on one hundred and eighty-two American-style options. Details on trades and volume of for every instrument are displayed in Table 6. The second one, denoted as “style”, includes American and European-style option trades on seventeen of these companies and covers a timespan going from January 1st 2014 to January 1st 2017. Indeed, since the introduction of European-style options on twenty-eight individual stocks in 2011, Eurex is the only exchange where traders can contemporaneously invest in American and European-style options on the same individual stock. As shown in Table 7, I collected European options data for all the firms, but American options data only for these seventeen companies (in bold). Nonetheless, in terms of European option trades, the dataset covers 96% of the market, permitting to spot almost all the manifestations of negative early exercise premia. These manifestations reflect either an arbitrage violation or an immediate gain of a LOAA trade.

4.2 Arbitrage violations and limit order approximate arbitrages

The simplest procedure to identify arbitrage violations would be based on intraday market quotes. Unfortunately, the Eurex data center does not provide this type of information for options. Nevertheless, together with the LOAAs, I can still infer such violations from trades, according to the methodology described as follows. To filter out pure arbitrage and limit order arbitrage trades involving a pair of options H and L , I first individuate, for both the samples, a set of market transactions characterized by:

- Synchronization: in the first/second sample at least two trades with the same size and with two different strikes/styles, occur at less than three seconds time difference.
- Immediate positive gain: trade price of L higher than the one of H .

Then, limit order arbitrage trades represent pairs of transactions featuring the same aggressor side, while pure arbitrage trades are executed on opposite aggressor sides (sell for L and buy for H). As a matter of fact, a trade where the aggressor is the seller indicates that the trade price was the Bid price right before the transaction. The same reasoning applies to the case where the aggressor is the buyer, if the trade occurred at the Ask price. This last criterion permits to collect a subset of trades involving a purchase at the Ask price of H and a sell at higher Bid price of L . The transactions are contemporaneous and produce a riskless profit, strongly suggesting that some market participants exploit the law-of-one price violations.

Relying only on market transactions, the procedure is able to pinpoint only violations that are exploited by some traders. There might have been negative free lunches in the order books, not taken from arbitrageurs. Under this perspective, the methodology is clearly a conservative estimate of the mispricings, not subject to the overestimation issues broadly documented in the literature, for instance by [Kamara and Miller \(1995\)](#) and [Chung \(1991\)](#).

To develop some intuition on the methodology, six pairs of synchronized transactions are exhibited in [Figure 7](#). At this stage, the data is already gathered and filtered according to the synchronization criterion, but LOAAs and PAOs must be still identi-

fied. The first and fifth pairs satisfy both the opposite aggressor side and immediate positive gain criterion and are consequently deemed as PAOs. The following three pairs incorporate a positive immediate gain and are traded at the same side. These are typical instances of successful LOAAs. On the contrary, the last pair is discarded, since, consistently with the law-of-one-price, entails a positive difference.

Table 3 displays the limit order and pure arbitrage trades for the two samples. LOAA trades are mainly implemented on options with different strikes. The reason is quite intuitive: unlike the strike strategy, the styles strategy embodies lower profits since at the option expiration the two payoffs cancel each other out. To clarify this point, I briefly describe their functioning: the arbitrageur submits a limit order on one of the two options and, in case of success, writes a European option and buys a less expensive American option on the same asset, thereby collecting the immediate gain at time t_M . The American option is not exercised until expiration T , when the total payoff of the strategy is null, independently of the price of the underlying.

PAOs are found for both the cases, proving not only that violations emerge in the order books of the two Eurex exchanges, but also that algorithmic trading strategies correct the mispricings, in line with the results of Akram et al. (2008) and Aiba and Hatano (2004). In the style strategies, violations produce negative early exercise premia. Their presence in option markets has been already reported by McMurray and Yadav (2000), who analyzed market quotes of FTSE100 options, and Dueker and Miller (2003), where the object of study is a 2-months period in 1986, when American and European-style options on the S&P500 were concurrently traded. Nonetheless, in both articles, the early exercise premium is computed using quotes, with arbitrage opportunities disappearing once Bid-Ask spreads are taken into account. Besides, the structure of their data does not permit to identify a scale-second quote, leading to possible difficulties in identifying synchronized prices for each of options styles. Conversely, the key advantage of this study is the data set, consisting in contemporaneous trades, which allows me to distinctly recognize the arbitrage operations instead of just revealing theoretical opportunities.

Strategy	Calls		Puts	
	Volume	Average return	Volume	Average return
Pure (Strike)	1453	299.78	192	1698.99
Pure (Style)	144	103.11	0	0
Approximate (Strike)	5753	534.22	1820	4547.30
Approximate (Style)	215	36.33	0	0

Table 3: Arbitrage trades. The volume of pure and limit order arbitrage trades in the two samples, together with the percentage average return. Pure arbitrages occur at opposite sides of the order book, whereas approximate arbitrages are implemented at the same side. The average return is computed by taking the limit and the market order fees as the total cost of the strategy.

4.3 Who are the arbitrageurs?

A natural question arises from examining Eurex options data: which category of market participant is able to construct a profitable *pure arbitrage trade*? I try to find a credible answer by linking the structure of order book fees to the results on pure arbitrage style trades.

Eurex allows to create three types of accounts: agent (brokers), proprietary (dealers) and market-makers. Within market makers, it further differentiates among regular (RMM), permanent (PMM) and advanced (AMM). Each one of these five categories can benefit from a reduced fee per order. Discount policies depend on whether the order is issued by a broker/dealer or by a market maker. Brokers and dealers can benefit from a reduce fee when the order exceeds 1000 and 500 contracts respectively. Market makers are instead refunded with monthly rebates, computed as percentages of the standard fee, once they satisfy certain requirements, with the most stringent ones for AMM. In particular, the single rebate corresponds to 55% for RMM and PMM, and to 80% for AMM. Table 4 summarizes the order book fees. The transaction costs are clearly decreasing with the degree of commitment of the participant to guarantee market liquidity. Dealers incur smaller fees than brokers, and market makers might benefit from the rebate system to partially offset their operating costs.

The differentiation in order fees allows, to some extent, to recognize the type of trader who is capable to obtain a net profit on style pure arbitrages. The overall costs and profits of this trade are realized at the same time. The arbitrageur invests in this strategy only if profits are higher than costs, with no uncertainty affecting the

decision. Since each type of market participant pays a specific fee, while profits are the same for each of them, costs are not. A pure style arbitrage trade could be potentially profitable for a market maker but not for a dealer, or, for a dealer, but not for a broker. Hence, combining the information from the arbitrage trades with the market fee structure, I am able to disentangle the fraction of arbitrages that could have been profitable only for a specific group of market participants. The approach is rather intuitive. First, I assume that brokers and dealers trade at a standard fee, since all the spotted arbitrage trades involve a number of contracts which is far below 500. By contrast, market makers always pay the reduced fee, serving the standard fee only as a base fee for the rebate calculation. Second, I specify a fee threshold corresponding to twice the amount paid for a single order, because an arbitrage on early exercise premium violations involve two operations: writing a European option and purchasing an American one. Afterwards, for each arbitrage trade, I compare the negative exercise risk premium with the threshold: if the proceeds are higher than the threshold then the arbitrage trade is profitable for a certain market participant. Brokers could have implemented only 9% of the arbitrage strategies with a positive gain, whereas market makers could have turned almost 90% of these trades into net profits.

Participant	Standard fee	Reduced fee	Fee threshold	% of arbitrages
Broker	0.15	0.080	0.30	0.09
Dealer	0.10	0.050	0.20	0.30
RMM and PMM	0.10	0.045	0.09	0.65
AMM	0.10	0.020	0.04	0.88

Table 4: Order book fees and arbitrages. The first two columns report the net order book standard and reduced fees. Brokers/dealers pay the reduced fee if the contract size is above 500/1000. Market makers never incur in standard fees, which represent only the base for computing the rebates, equivalent to 55% for Regular (RMM) and Permanent (PMM) Market Makers, and to 80% for Advanced Market Makers (AMM). The fee threshold is the cost to implement an arbitrage per single contract, equal to twice the order book fee. For brokers/dealers the fee threshold is computed using the standard fee, since in the sample there are no arbitrage trades involving transactions with more than 500/1000 contracts. The percentage (%) of arbitrages is then computed as the the number of negative early exercise premia smaller than the threshold and the total number of negative early exercise premia.

The small size of the sample motivates a non-parametric bootstrap analysis. I re-sample the outcomes of the negative early exercise premium 10000 times to obtain

a kernel density estimation of its distribution, illustrated in Figure 8. The shaded polygons can be interpreted as profitable areas, delimited by the fee thresholds on the vertical straight lines. The most substantial gains (the green polygon) can be obtained by all the market participants, but represent only a tiny portion of the whole set of arbitrages. In fact, a sizeable part of them (the red and orange polygons together, corresponding to 70% of the total area) is accessible only to market makers.

5 Conclusions

I have introduced the notion of limit order approximate arbitrages and explained its functioning in a framework incorporating option pricing models and order book dynamics. Moreover, I documented the existence of these trades in the Eurex exchange, one of the most liquid existing option markets. I have also clarified that, occurring at the same side of the order book (Bid or Ask) for both assets, a limit order approximate arbitrage trade is slightly different from a pure arbitrage trade. Unlike the latter one, it does not in fact stem from an arbitrage violation.

Being its gain-loss ratio extremely large, a LOAA constitutes an unusually attractive opportunity for an investor. The initial expenditure, coinciding with the order fee linked to the aggressive limit order, can be rewarded by both an immediate positive gain and a vertical spread option payoff. As already specified, one option should ordinarily be more expensive than the other one, due to the different strikes or styles of the two assets. Nevertheless, at some random point in time, one of the two Bid-Ask spreads becomes large enough to allow the employment of the LOAA. Eventually, an investor exploits the abnormal width of one of the two spreads by placing an aggressive limit order to construct the LOAA position.

The model demonstrates how, regardless of the values assigned to its parameters, an agent with logarithmic utility engages in the strategy in almost all market conditions. She decides not to submit the aggressive limit order only when abnormal parameter values are plugged-in.

In addition, this paper enriches the literature on arbitrage violations, by identifying pure arbitrage trades on both strikes and styles. In the case of options with different styles, these are novel findings. Indeed, taking advantage of an almost unique pecu-

liarity of the Eurex market, I directly measure the early exercise premium and prove that a remarkable fraction of the European option market trades are triggered when negative early exercise premia are observed.

Extensions to this study might consist in examining limit order approximate arbitrages for other assets such as stocks, for example by designing strategies built on the put-call parity relation, and foreign exchange rates, for instance by exploiting the differences among spreads of different currencies. Exploring the profitability of the trading scheme described by this paper in other contexts would help to increasingly fill the gap between the theory on almost riskless opportunities and its practical implementation in financial markets.

References

- Aiba, Y. and N. Hatano (2004). Triangular arbitrage in the foreign exchange market. *Physica A: Statistical Mechanics and its Applications* 344(1), 174–177.
- Ait-Sahalia, Y., R. Kimmel, et al. (2007). Maximum likelihood estimation of stochastic volatility models. *Journal of financial economics* 83(2), 413–452.
- Akram, Q. F., D. Rime, and L. Sarno (2008). Arbitrage in the foreign exchange market: Turning on the microscope. *Journal of International Economics* 76(2), 237–253.
- Bernardo, A. E. and O. Ledoit (2000). Gain, loss, and asset pricing. *Journal of political economy* 108(1), 144–172.
- Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy* 81(3), 637–654.
- Bondarenko, O. (2003). Statistical arbitrage and securities prices. *The Review of Financial Studies* 16(3), 875–919.
- Chung, Y. P. (1991). A transactions data test of stock index futures market efficiency and index arbitrage profitability. *The Journal of Finance* 46(5), 1791–1809.
- Cochrane, J. H. and J. Saa-Requejo (2000). Beyond arbitrage: Good-deal asset price bounds in incomplete markets. *Journal of Political Economy* 108(1), 79–119.

- Cont, R., S. Stoikov, and R. Talreja (2010). A stochastic model for order book dynamics. *Operations research* 58(3), 549–563.
- Dueker, M. and T. W. Miller (2003). Directly measuring early exercise premiums using american and european s&p 500 index options. *Journal of Futures Markets* 23(3), 287–313.
- Evnine, J. and A. Rudd (1985). Index options: The early evidence. *The Journal of Finance* 40(3), 743–756.
- Foucault, T., O. Kadan, and E. Kandel (2005). Limit order book as a market for liquidity. *Review of Financial Studies* 18(4), 1171–1217.
- Foucault, T., R. Kozhan, and W. W. Tham (2017). Toxic arbitrage. *Review of Financial Studies* 30(4), 1053–1094.
- Gould, J. P. and D. Galai (1974). Transactions costs and the relationship between put and call prices. *Journal of Financial Economics* 1(2), 105–129.
- Gromb, D. and D. Vayanos (2010). Limits of arbitrage. *Annu. Rev. Financ. Econ.* 2(1), 251–275.
- Grossman, S. J. and J. E. Stiglitz (1976). Information and competitive price systems. *The American Economic Review* 66(2), 246–253.
- Grossman, S. J. and J. E. Stiglitz (1980). On the impossibility of informationally efficient markets. *The American Economic Review* 70(3), 393–408.
- Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies* 6(2), 327–343.
- Kamara, A. and T. W. Miller (1995). Daily and intradaily tests of european put-call parity. *Journal of Financial and Quantitative Analysis* 30(04), 519–539.
- MacKinlay, A. C. and K. Ramaswamy (1988). Index-futures arbitrage and the behavior of stock index futures prices. *Review of Financial Studies* 1(2), 137–158.

- McMurray, L. and P. K. Yadav (2000). The early exercise premium in american option prices: Direct empirical evidence.
- Muravyev, D., N. D. Pearson, and J. P. Broussard (2013). Is there price discovery in equity options? *Journal of Financial Economics* 107(2), 259–283.
- Parlour, C. A. (1998). Price dynamics in limit order markets. *Review of Financial Studies* 11(4), 789–816.
- Peña, I., G. Rubio, and G. Serna (2001). Smiles, bid-ask spreads and option pricing. *European Financial Management* 7(3), 351–374.
- Rhee, G. S. and rosita p. chang.(1992). intraday arbitrage opportunities in foreign exchange and eurocurrency markets. *The Journal of Finance*, 363–379.
- Roşu, I. (2009). A dynamic model of the limit order book. *Review of Financial Studies* 22(11), 4601–4641.
- Shanken, J. (1992). The current state of the arbitrage pricing theory. *The Journal of Finance* 47(4), 1569–1574.
- Shleifer, A. and R. W. Vishny (1997). The limits of arbitrage. *The Journal of Finance* 52(1), 35–55.
- Stoll, H. R. (1969). The relationship between put and call option prices. *The Journal of Finance* 24(5), 801–824.
- Stoll, H. R. and R. E. Whaley (1990). The dynamics of stock index and stock index futures returns. *Journal of Financial and Quantitative analysis* 25(04), 441–468.

A Arbitrage at the Bid

At time t_M the agent observes the order book in Table 5. At t_{M+1} she submits a buy order in-between the best two Bids, realizing, in case of an immediate buy market order for $C_{100}(t_M, T)$, a profit equal to $1.3 - 1.2$ together with the payoff of the vertical spread BCS_T .

$C_{100}(t_M, T)$		$C_{110}(t_M, T)$		$C_{100}(t_{M+1}, T)$		$C_{110}(t_{M+1}, T)$	
Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask
0.8		1.2		0.8		1.2	
1.0		1.3		1.0		1.3	
	1.5		1.4	1.2			1.4
			1.5		1.5		1.5

Table 5: Strategy at the Bid price. An example of the strategy when the Bid-Ask spread of $C_{110}(t_M, T)$ is included in the one of $C_{100}(t_M, T)$. The agent aggressive limit order is highlighted in bold.

B Proofs

In this section I provide the proof of Proposition 2.3.

Proof. Given an option $O = \{H, L\}$, let M_O be the time until the first buy market order for the option arrives, L_O the time until a sell limit order enters as best Ask, and T_O the time until the first cancellation order for the best Ask arrives. Then:

$$\begin{aligned}
p_S &= \int_0^\infty P(M_O > t | t = X) P(T_O > t | t = X) P(L_O = t) dt \\
&= \int_0^\infty P(M_O > t) P(T_O > t) P(L_O = t) dt \\
&= \int_0^\infty e^{-\mu_O t} e^{-\gamma_O t} \lambda_O e^{-\lambda_O t} dt = \frac{\lambda_O}{\mu_O + \gamma_O}.
\end{aligned} \tag{14}$$

The first equation is simply the definition of the probability of one of three events happening before a fourth one. In the second equation the conditioning is removed, being the events mutually independent. Poisson densities and cumulative distribution functions are plugged-in the last equation where the integral is then easily computed.

□

C Figures

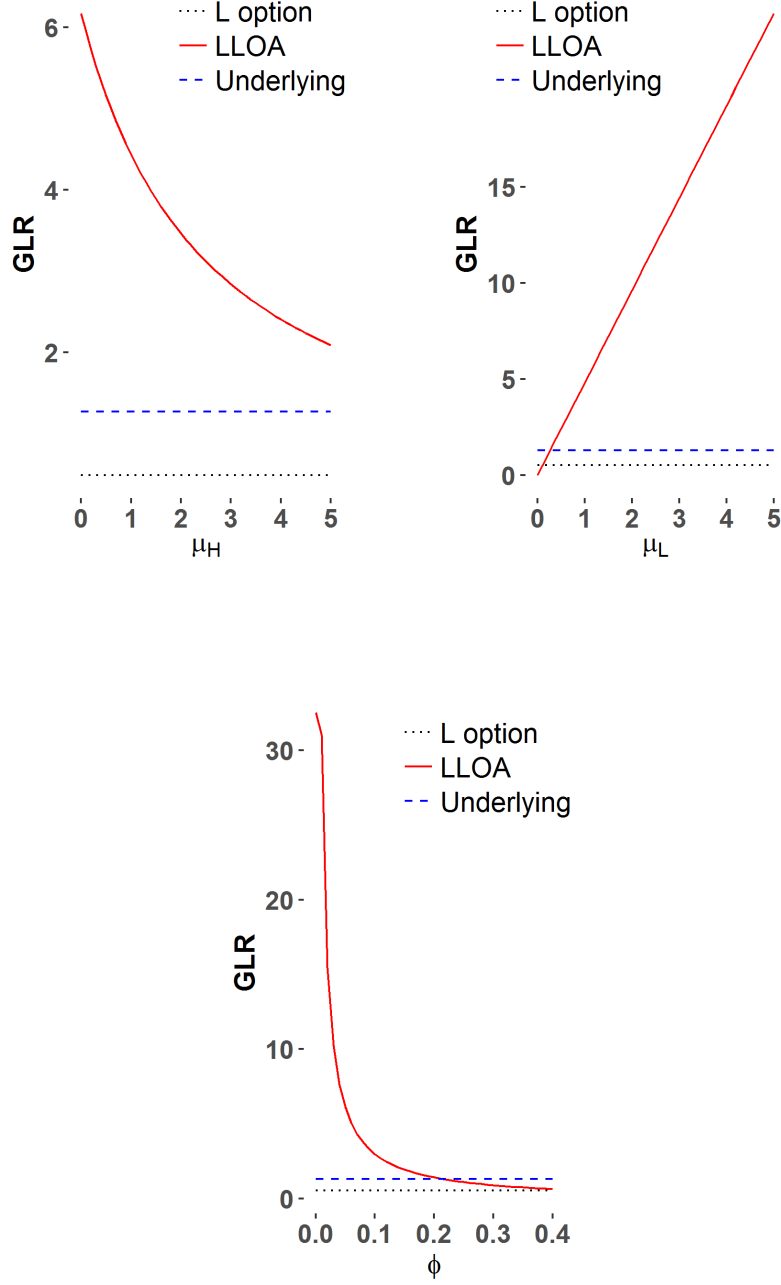


Figure 2: Gain-loss ratio in a Black Scholes setting. The gain-loss ratio $GLR := E[\{x, 0\}^+]/E[\{-x, 0\}^+]$ is computed for the underlying S_t , for the option L and the for the LOAA. The S_t dynamics is given by $dS_t = \eta S_t dt + \sigma S_t dB_t$, with a Brownian motion B_t , $\eta = 0.02$ and $\sigma = 0.2$. The other parameter values are $\mu_L = 0.94$, $\mu_H = 0.94$, $\lambda_L = 1.95$, $\gamma_L = 0.71$, $\phi = 0.065$, $S_t = 90$ and $T = 6/12$.

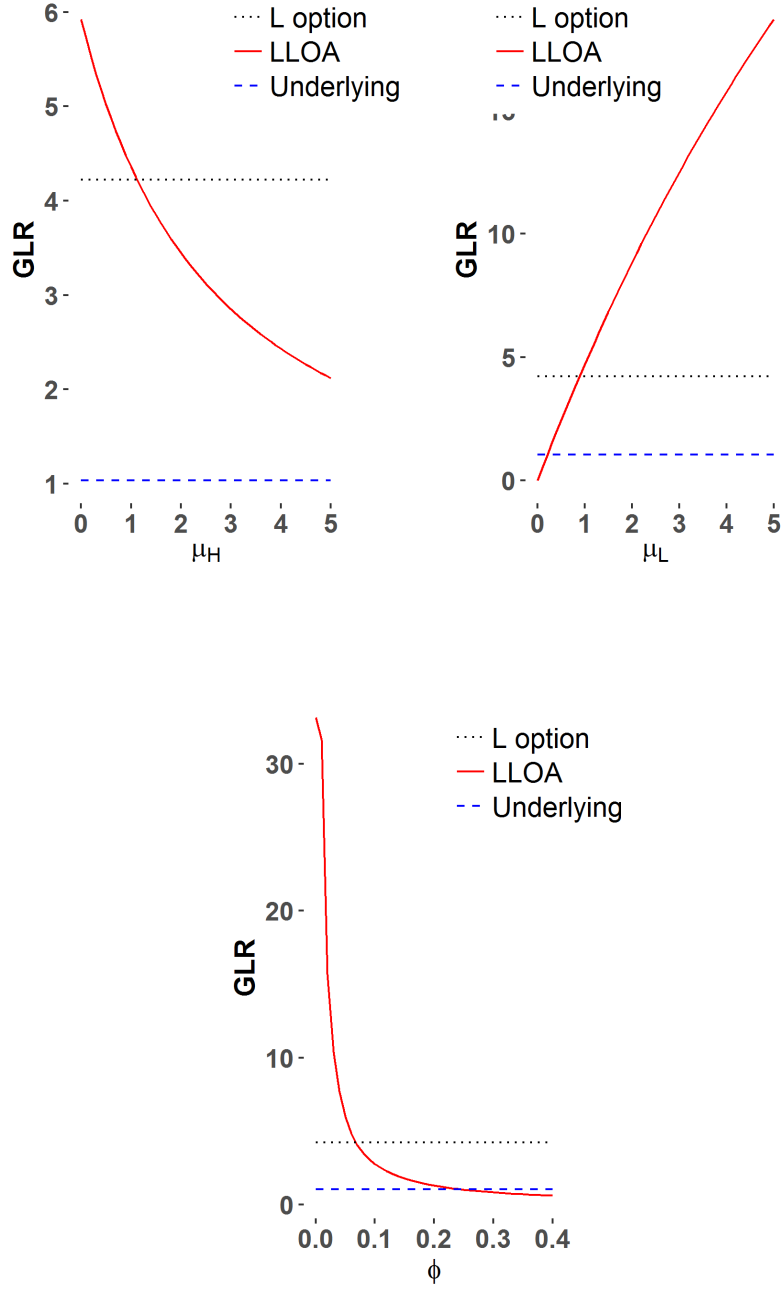


Figure 3: Gain-loss ratio in a Heston setting. The gain-loss ratio $GLR := E[\{x, 0\}^+]/E[\{-x, 0\}^+]$ is computed for the underlying S_t , for the option L and the for the LOAA. The S_t dynamics is given by $dS_t = \eta S_t dt + \sqrt{v_t} dB_t^1$ and $dv_t = \kappa(\alpha - v_t)dt + \epsilon dB_t^2$, where B_t^1 and B_t^2 , are two Standard Brownian motions with $Cov(dB_t^1, dB_t^2) = \rho dt$, $\eta = 0.02$, $\rho = -0.767$, $\epsilon = 0.048$, $\kappa = 5.07$, $v_{t_M} = \sqrt{0.2}$, $\alpha = 0.0457$. The other parameter values are $\mu_L = 0.94$, $\mu_H = 0.94$, $\lambda_L = 1.95$, $\gamma_L = 0.71$, $\phi = 0.065$, $S_t = 90$ and $T = 6/12$.

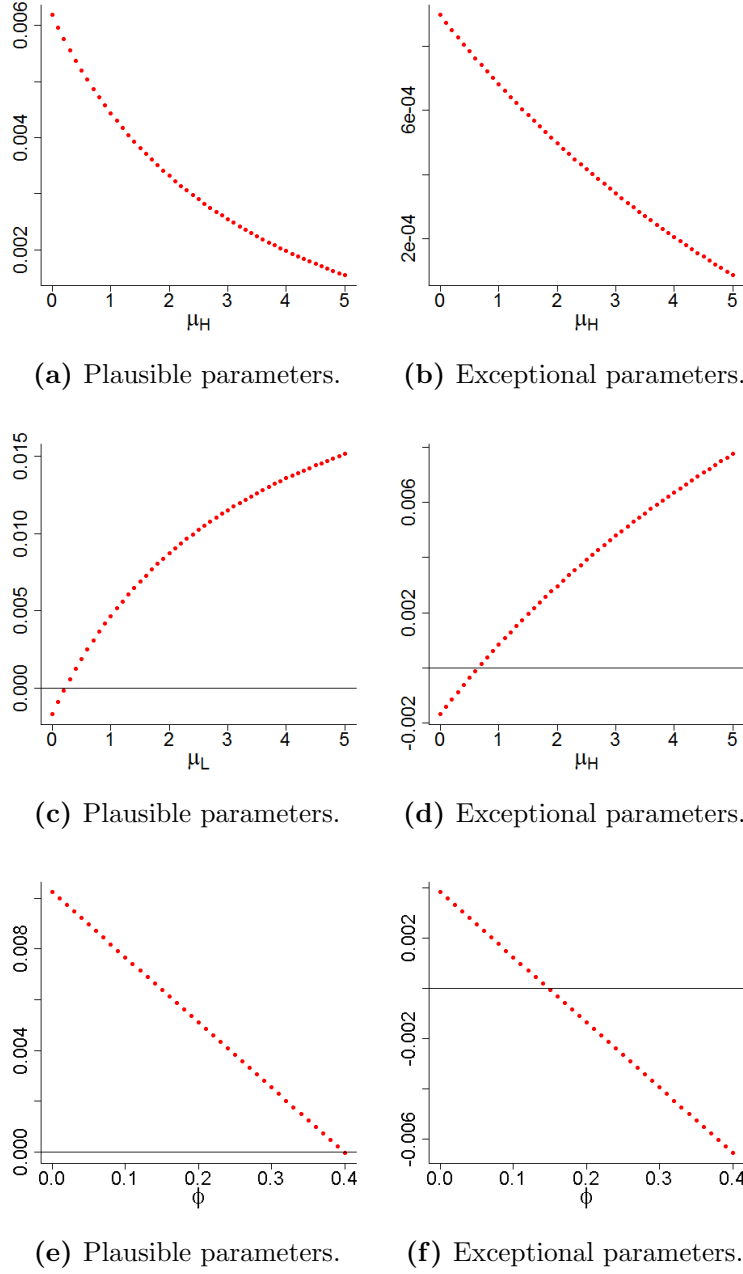


Figure 4: $E_{t_M}(\log(100 - SP)) - \log(W)$ in a Black-Scholes setting. The S_t dynamics is given by $dS_t = \eta S_t dt + \sigma S_t dB_t$, with a Brownian motion B_t . The fixed variables are $W = 100$, $S_{t_M} = 90$, $T = 6/12$. The parameter values are $\eta = 0.02$, $\sigma = 0.2$, $\mu_L = 0.94$, $\mu_H = 0.94$ and $\phi = 0.065$. The plausible values for the other parameters are $\lambda_L = 1.95$ and $\gamma_L = 0.71$. The exceptional ones are $\lambda_L = 5$ and $\gamma_L = 5$.

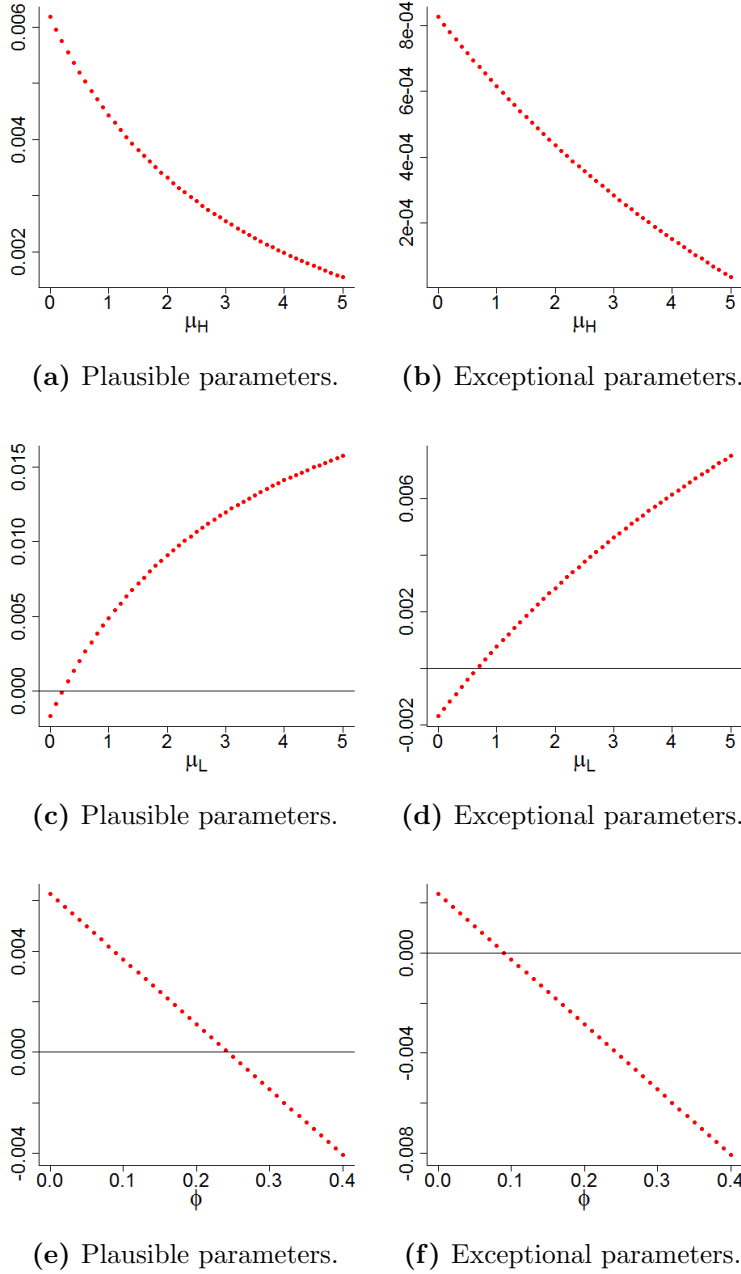
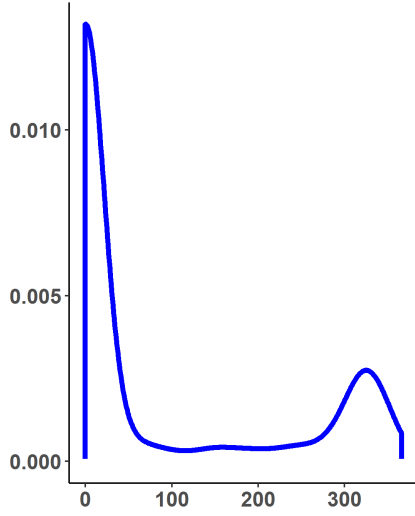
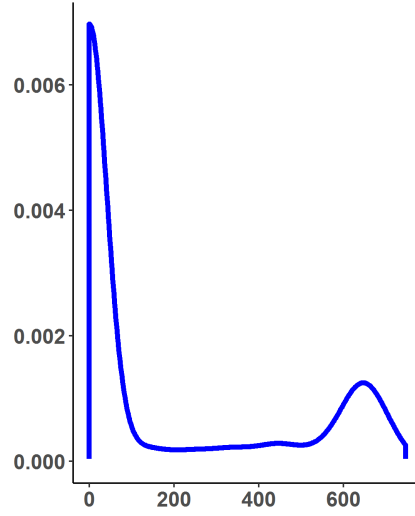


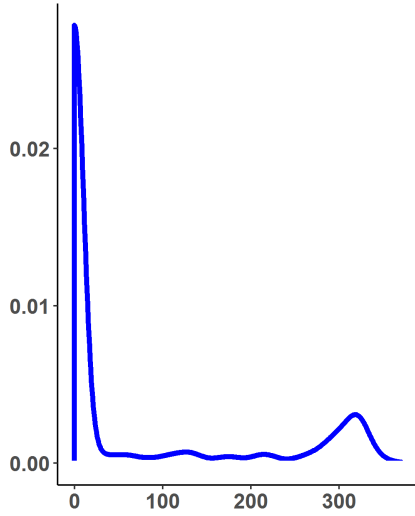
Figure 5: $E_{t_M}(\log(100 - SP)) - \log(W)$ in a Heston setting. The dynamics of the underlying S_t is modeled as $dS_t = \eta S_t dt + \sqrt{v_t} dB_t^1$ and the one of v_t is described by $dv_t = \kappa(\alpha - v_t)dt + \epsilon dB_t^2$, where B_t^1 and B_t^2 are two Standard Brownian motions with $Cov(dB_t^1, dB_t^2) = \rho dt$. The fixed variables are $W = 100$, $S_{t_M} = 90$, $T = 6/12$. The parameters values are $\eta = 0.02$, $\rho = -0.767$, $\epsilon = 0.048$, $\kappa = 5.07$, $v_{t_M} = \sqrt{0.2}$, $\alpha = 0.0457$, $\mu_L = 0.94$, $\mu_H = 0.94$ and $\phi = 0.065$. The plausible values for the other parameters are $\lambda_L = 1.95$ and $\gamma_L = 0.71$, while the exceptional ones are $\lambda_L = 5$ and $\gamma_L = 5$.



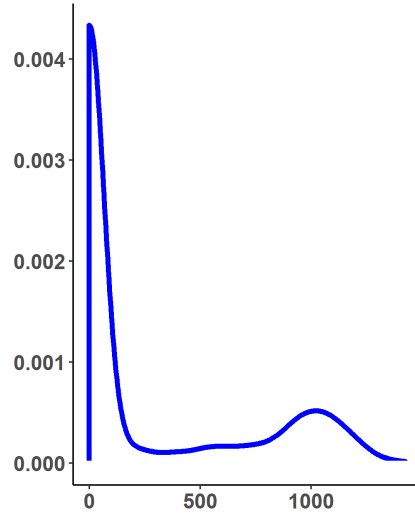
(a) Heterogeneous spreads, $\delta = 0.01$.



(b) Heterogeneous spreads, $\delta = 0.005$.



(c) Homogeneous spreads, $\delta = 0.01$.



(d) Homogeneous spreads, $\delta = 0.005$.

Figure 6: Return distribution. A kernel-density of the LOAA return distribution is plotted. Given $T = 6/12$, the prices $C_{100}^{Bid}(t, T)$, $C_{100}^{Ask}(t, T)$, $C_{110}^{Bid}(t, T)$ and $C_{110}^{Ask}(t, T)$ are simulated at every minute. Considering a day of seven hours and a month of 30 days, the total number of observation is equal to 75600. The underlying is simulated according to a GBM with $\mu = 0.02$, $\sigma = 0.2$ and starting value $S_t = 90$. The fundamental value is computed with the Black-Scholes formula. Fixing $\lambda_L = 1$ and $\mu_H = 1$, spreads are homogeneous when $\lambda_H = \mu_L = 1$ and heterogeneous when $\lambda_H = \mu_L = 0.8$. The other parameters are $\phi = 0.065$ and $\gamma_H = \gamma_L = 0.4$.

SecurityID	MDEntryTime	Price	Aggr. side	Size	Style	Strike	Type	Expiry	Underlying
382624	09:24:33.792495966	8.05	Sell	1	E	120	Call	6/19/2015	BAY
382606	09:24:33.806656830	7.91	Buy	1	A	120	Call	6/19/2015	BAY
9337	08:18:34.297101124	0.66	Buy	50	A	120	Put	6/19/2015	ALQ
9341	08:18:34.295143782	0.9	Sell	50	A	110	Put	6/19/2015	ALQ
15200	12:10:56.731642365	4.06	Sell	7	A	170	Put	6/19/2015	LVM
15253	12:10:56.732547662	4.66	Sell	7	A	155	Put	6/19/2015	LVM
119	08:05:02.181006063	3.3	Sell	4	E	100	Call	12/16/2016	NOV
36766	08:05:02.995632898	2.85	Sell	4	A	100	Call	12/16/2016	NOV
77190	08:09:10.137141319	0.27	Buy	20	A	160	Call	02/20/2015	BAY
77190	08:09:10.136059623	0.2	Buy	20	A	155	Call	02/20/2015	BAY
98868	09:09:46.950122819	2.87	Buy	1	E	200	Call	12/16/2016	BAY
98869	09:09:46.950122819	2.9	Buy	1	A	200	Call	12/16/2016	BAY

Figure 7: Synchronized transactions. The figure shows six examples of synchronized transactions. The columns report security id, trade time and price, aggressor side, size, option-style, strike price, option-type, expiration and underlying stock. The individual companies appearing in this example are Bayer (BAY), L’Oreal (LVM), Novartis (NOV) and Air Liquide (ALQ). Pure arbitrages and limit order arbitrages are in dark green and green respectively. The last pair (in red) is not included in any of the two groups.

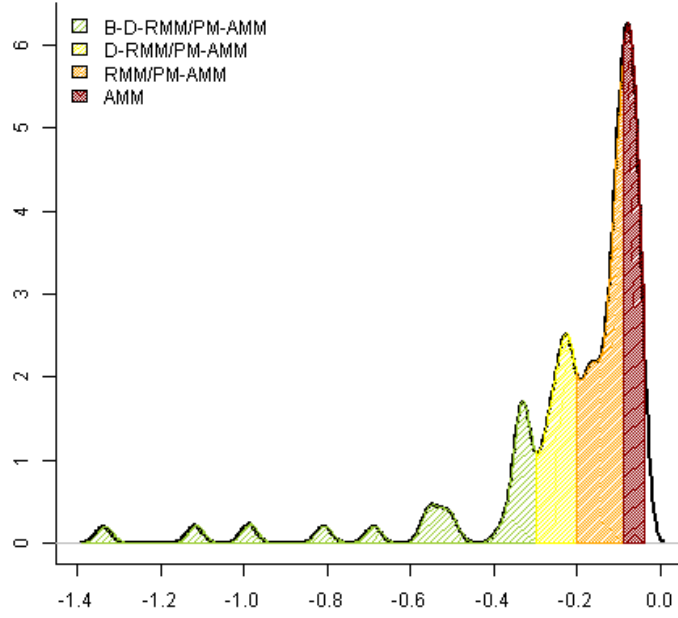


Figure 8: Negative early exercise premium distribution. The chart shows the bootstrapped distribution of the negative early exercise premium. The bootstrap is implemented by constructing 10000 re-samples without replacement. The kernel density estimation is given by the black line. The green polygon represents the fraction of arbitrages accessible to all the market participants (broker, dealers, market makers) and the red polygon represents the arbitrages profitable only to Advanced Market Makers (AMM). Intermediate cases are in-between these two areas.

D Dataset

COMPANY	TRADES	VOLUME	COMPANY	TRADES	VOLUME
Aalberts	616	8346	Hugo Boss	8097	105139
Aareal Bank	3706	6186	Immobilien	2116	271761
ABN AMRO	948	50658	Inditex	15190	344207
Accor	8109	85860	Infineon	71027	4709361
Ackermans	85	463	Intesa	16468	779881
Actellion	34790	507758	Jenoptix	702	7196
Adecco	38676	643240	Julius Baer	15404	331229
Adidas	164769	1678910	Kardex	80	532
Aegon	13328	1186301	Klepierre	202	1639
Ahold	8513	230608	Kloeckner & co	10107	317839
Air France	10386	447514	Koninklijke	15463	190150
Air Liquide	33223	423123	Kontron	116	3160
Airbus	6312	17285	Kuehne & Nagel	11567	112207
Aixtron	13005	619468	Kuka	2870	26947
Akzo Nobel	13758	139457	Lafarge	4241	81118
Alcatel	67	5910	Lagardere	1983	36170
Allianz	263702	3578767	Lenzing	678	8063
Alpiq	29	1620	Leoni	6217	80476
AMS	6900	127467	Lindt & Spruengli	1862	4755
Anglo American	1431	8011	Lindt PS	7255	72861
Anheuser-Busch	34853	531614	Lloyds	1506	76392
Aperam	4758	45774	Lonza	22109	180107
Arytza	12172	186214	L'Oreal	38456	363249
ASML Holding	29062	308414	LPKF	42	964
Aurubis	6339	78103	Lukoil	338	19473
AXA	59904	2453561	Luxottica	668	5793
Axel Springer	882	14773	LVMH	54241	527126
BAE Systems	353	4426	Mediobanca	673	9508
Baloise Holding	17888	176776	Meyer Burger	5037	165381
Banco Bilbao	15509	1354492	Michelin	15670	149034
Banco Santander	23539	2305931	Muenchener	89604	914156
Barclays	1095	36350	Nestlé	143921	8437201
BASF	354386	5985441	Nestlé Oil	4846	60487
Bayer	324786	4923600	Novartis	139901	5408042
Baywa	841	9807	Oerlikon	1560	20617
BB Biotech	1370	10438	Ost Post	8374	435928
Beiersdorf	40345	377061	Pernod-Ricard	9185	140266
Bilfinger	21229	253701	Peugeot	9656	532014
BMW	256748	4426005	Pirelli	237	3401
BNP Paribas	52208	1442904	Porsche	18312	240719
Boskalis	4328	53256	Prosiebensat	8384	147506
Bouygues	14309	293378	Publicis Groupe	4846	56056
Capgemini	14659	227323	Raiffeisen	6537	139488
Carrefour	18974	674573	Rational	263	1497
Casino Guichard	11543	116762	RDX Index	228	12323
Celesio	810	26529	Reed Elsevier	2341	45957
Clariant	24666	1225718	RELX	2344	67655
Continental	93037	789535	Renault	33277	443602
Coomerzbank	137272	9567598	Rheinmetall	7864	94519

Corio	1020	10905	RHI	946	31266
Credit Suisse	115284	7440966	Rio Tinto	1407	6986
Daimler	426653	10077573	Rocket Internet	1642	27675
Danon	32780	568581	Ryanair	266	7978
Delhaize	4947	46487	Safran	9099	135748
Delta Lloyd	4868	128033	Saint Gobain	23498	525569
Deutsche Bank	377745	15195967	Salzgitter	21323	297214
Deutsche Boerse	35574	554511	Schneider	28972	504701
Deutsche shop	1241	17940	SGL Carbon	10273	185326
Deutsche Post	148186	6511816	SGS	23385	176469
Deutsche Telekom	141756	13404812	Sika	4491	91786
Deutsche Wohnen	792	28372	Sky	41	510
Dialog	4126	52370	SNAM	2264	37705
DKSH	2090	35911	Sodexo	3072	27353
Drillisch	1859	21833	Software	2104	36088
DT.Annington	72	810	Sonova	17212	146180
Dufry	5273	61773	STMicroel.	5444	471138
EdF	10673	226717	Stora Enso	2074	37368
Enel	11630	559003	Suedzucker	9822	216670
ENI	16935	447288	Swatch	13322	131248
Ericcson	501	10442	Swiss Life	26567	235476
Erste	2559	43905	Technip	14110	159583
Essilor	20655	194930	Telecom Italia	8241	551133
Eurazeo	690	22783	Telekom Austria	440	62183
Evonik	6400	154838	Tom Tom	2523	104904
Fielmann	2454	25048	Transocean	35512	1030852
Fortum	7199	171881	TUI	17029	530688
Freenet	7488	167991	Unibail-Rodamco	21516	130303
Fresenius	60565	642699	Unicredit	5861	84809
Fuchs	1931	24657	United Internet	809	12947
Fugro	5823	72925	UPM-Kymmene	2957	93293
Gazprom	1371	102850	Valeo	9635	217556
GEA	6048	103353	Veolia	12103	362211
Geberit	20435	126012	Verbund	2092	37992
Generali	20285	1236325	Vinci	27550	481655
Givaudan	37309	291831	Vivendi	14215	633116
Glencore	4047	74854	Vodafone	672	43456
Ham. Hafen	583	16866	Voestalpine	3090	49784
Hannover Rueck	20468	169169	Wendel	233	9392
Heineken	11274	122630	Wincor	6865	78196
Hochtief	3724	35627	Wirecard	14150	292778
HSBC	1490	30190			

Table 6: Strike sample: Eurex American options. The table reports the total number and volume of trades for each option on individual stock. The dataset includes all the transactions registered in the Eurex option market from January 1st 2014 to May 23th 2017.

COMPANY	TRADES		VOLUME	
	American	European	American	European
ABB		23		1687
Adidas	147,023	2	1,537,087	4
Allianz	212,476	20	6,462,610	274
AXA	197,057	24	3,884,222	420
BASF	316,001	15	5,703,757	177
Bayer	297,417	1,029	4,777,220	6,114
BMW	228,139	1	4,183,093	13
Credit Suisse		148		5060
Commerzbank	88,399	120	6,604,778	845
Daimler	244,056	9	6,524,401	280
Deutsche Bank	326,823	16	14,468,560	984
Deutsche Boerse	35,574	1	554511	200
Deutsche Telekom	126,362	9	12,767,154	362
E.ON	5,548	40	522,811	2207
Infineon	69328	1,229	411352	9,066
K+N		1		120
Metro		1		25
Muenchener Rueckversiche	82301	96	890339	369
Nestlé	134290	158	7900338	1165
Nokia		53		4,391
Novartis	134677	366	5117882	3,495
Oracle		19		726
Roche		13		772
RWE		6		55
Siemens		12		2,888
Swiss Re		69		1,677
VW		12		56
Zurich		50		1,008
Total		3312		39892

Table 7: Style sample: Eurex American and European-style options. The table reports the total number and volume of trades for each option on individual stock where both the styles are traded. The dataset includes all the transactions registered in the Eurex option market from January 1st 2014 to January 19th 2017.

Trading central moments^{*}.

Paul Schneider [†] Davide Tedeschini^{*‡}

Abstract

Mean, variance, and skewness of the return of an asset are important measures for our economic understanding. We propose a definition of *realized central moments* that is tradeable. The prices of these realized central moments are the implied central moments proposed by [Bakshi et al. \(2003\)](#). The motivation for utilizing our measure rather than sample moments is threefold: first, unlike sample moments, our measures are tradeable and the trading profits therefore admit an interpretation as risk premia. Second, even if sample central moments were tradeable, asymptotically, they could be different from implied central moments in absence of risk premia. Finally, estimates of sample central moments are based on the past trajectory of the financial asset, while our measures, as well as implied moments, are not.

Evaluation of variance, skewness and kurtosis risk premia represents a crucial topic in the finance literature. Their definition always involves two components: the current price of a contract, with one of the above-mentioned measures as underlying, and its corresponding realized payoff. An influential article of [Bakshi et al. \(2003\)](#) (hereafter BKM) introduces a procedure to calculate in a model-free way the price of central moments. Their work takes full advantage of several previous findings on the estimation of implied measures based on the information contained in option prices. This strand of

^{*}We have benefited from helpful discussions on this topic with Gurdip Bakshi, Christopher Hemmens (discussant), Jerome Detemple (discussant), Dilip Madan, Roberto Marfe (discussant), Anthony Neuberger and Mirela Sandulescu. Participants at SoFIE Conference on Financial Econometrics & Empirical Asset Pricing 2016 (Lancaster), Swiss Doctoral Workshop 2016 and Finance Workshop 2016 (Turin) provided many useful suggestions. Financial support from Swiss National Science Foundation (projects “Trading Asset Pricing Models” and “Equity risk premia: model evaluation, trading strategies and estimation of implied moments”) is gratefully acknowledged.

[†]University of Lugano and Swiss Finance Institute.

[‡]University of Lugano, Swiss Finance Institute.

research can be divided into two subgroups, with the first one focusing on the estimation of the risk-neutral density, starting from the seminal paper of [Breedon and Litzenberger \(1978\)](#) and including, among others, most recent developments such as [Aït-Sahalia and Lo \(1998\)](#) and [Bakshi and Panayotov \(2010\)](#), and the second one aiming at inferring the moments of the distribution. The resulting option-based quantities were derived by [Britten-Jones and Neuberger \(2000\)](#) exploiting solely the information of call options and by [Carr and Madan \(2001\)](#), whose method requires both out of the money calls and puts. One of the most important applications of the latter work is the construction of an option-based measure for central moments, as shown by BKM. Their approach is broadly implemented not merely for research purposes, but also for designing indexes to be traded by practitioners. A remarkable example in this context is represented by the CBOE SKEW index, launched by the CBOE in order to produce a nonparametric proxy for the S&P500 log returns skewness.

For this reason, the tool for computing the expectation under the forward pricing measure of variance, skewness and kurtosis is fully available. Unfortunately, as previously described, the measurement of the risk premia related to central moments would not be accurate without setting up an equivalent realized quantity, a task not yet carried out in the literature. This paper intends to fill this gap by introducing n th moment swaps with BKM implied moments as fixed legs, and with floating legs satisfying certain characteristics. Essentially, we need the n th moment risk premium to be the expected profit from a trading strategy and to reflect exactly BKM prices. In absence of arbitrage opportunities the tradeability requirement plays a pivotal role, as already highlighted by [Kozhan et al. \(2013\)](#). Indeed, any fundamental theorem of asset pricing which puts the structure of no arbitrage on prices includes the price as an expectation of a pay-off which is traded directly or can be replicated. Thus, a definition of a risk premium compatible with the no-arbitrage assumption can only be as the expected profit of a trading strategy. As argued below, the structure of floating and fixed legs for central moments is intrinsically different, causing the related swap contract to be non negotiable in financial markets. Our approach to overcome the issue consists in centring both the pay-off and the price of the strategies with respect to the forward pricing measure. In this fashion, the variance, skewness and kurtosis contracts become perfectly replicable with forward and options.

In addition, we recover an approximation for the coefficient of relative risk aversion of a power-utility agent investing in the central moment swaps we constructed. In the same spirit of [Ant-Sahalia and Lo \(2000\)](#) and [Jackwerth \(2000\)](#), who recover utility functions from options-implied and physical distributions, our expression links relative risk aversion to their option-implied and realized moments.

Another important aspect we examine in the paper concerns the implications of adopting sample moments for computing risk premia. The investigation on the presence of risk premia is often conducted by comparing the sample moments of the physical distribution and the average implied moments of the forward pricing distribution. We demonstrate analytically that this scheme is legitimate for non-central moments but might lead to spurious results in the case of central moments. An empirical analysis provides evidence that the usage of sample moments as BKM implied moments counterparts should be strongly discouraged, seeing that the two quantities appear rather unconnected. In order to provide additional evidence of this finding, we make a comparison between an estimation of the coefficient of relative risk aversion involving sample moments and an analogue experiment exploiting our theoretical approximation. Results show that sample moments might entail negative values for the relative risk aversion, clearly violating the standard assumption of positive relative risk aversion.

The paper is organized according to the following outline. In [Section 1](#), the tradeable realized measures are presented together with a straightforward argument justifying their use. [Section 2](#) discusses the theoretical issue arising when comparing sample moments from the physical distribution and average forward pricing moments for the identification of risk premia. [Section 3](#) is devoted to the empirical analysis and [Section 4](#) concludes. Most of the proofs are included in the main body of the article but some specific cases are left for the Appendix, which also contains all the figures.

1 A tradeable measure for central moments

The goal of this section consists in defining the risk premia obtainable by trading central moments. To this end, we work in a single-asset economy where a forward contract on this risky asset is traded together with call and put options. The physical measure is denoted by P and the forward pricing measure Q_T is used in order to avoid irrelevant

issues related to the stochasticity of the interest rates. The forward price of the asset at time t is $F_{t,T}$, its log return is $R_{t,T} := \log(F_{T,T}/F_{t,T})$ and risk premia are defined as follows.

Definition 1.1. A risk premium is the P-expected pay-off from a trading strategy.

To appreciate the definition consider an investor willing to construct swaps that hedge the central moments. The conceptual problem can be seen in the case of a buy-and-hold strategy, where the unattainable pay-off at time t will be:

$$(R_{t,T} - E_t^P(R_{t,T}))^n - E_t^{Q_T} \left((R_{t,T} - E_t^{Q_T}(R_{t,T}))^n \right) \quad (1)$$

and the unattainable risk premium associated with this strategy is:

$$E_t^P \left((R_{t,T} - E_t^P(R_{t,T}))^n \right) - E_t^{Q_T} \left((R_{t,T} - E_t^{Q_T}(R_{t,T}))^n \right). \quad (2)$$

The pay-off is unattainable because the conditional P expectation is not known, and because the difference between the n th central moments under the P and Q_T measures is not the expected pay-off to a trading strategy (due to the centring under different measures) and so does not qualify as a risk premium. Knowing the n th central moments of the distribution under both measures and observing that they differ at some point in time does not provide a strategy for earning expected excess returns. Hence, it is not evidence of the existence of a risk premium consistent with Definition 1.1.

A reasonable solution to this issue consists in replacing the non-central moment $E_t^P(R_{t,T})$ with its Q_T equivalent $E_t^{Q_T}(R_{t,T})$ in Equation (1):

$$\left(R_{t,T} - E_t^{Q_T}(R_{t,T}) \right)^n - E_t^{Q_T} \left(\left(R_{t,T} - E_t^{Q_T}(R_{t,T}) \right)^n \right), \quad (3)$$

so that the realized leg of the swap is centered under the same measure and the strategy becomes fully tradeable. Indeed, the forward pricing t -conditional first moment $E_t^{Q_T} \left(\left(R_{t,T} - E_t^{Q_T}(R_{t,T}) \right)^n \right)$ is replicable by a portfolio of out of the money call and put options on $F_{t,T}$, as BKM have demonstrated. These risk-neutral central moments are widely used and here we present them again for completeness reasons, but also along with their realized counterparts $E_t^P \left(\left(R_{t,T} - E_t^{Q_T}(R_{t,T}) \right)^n \right)$. Consistent with BKM

notation we introduce the implied polynomial moments of log-forward returns:

$$\mu_{t,T} := E_t^{\mathbb{Q}^T}(R_{t,T}), V_{t,T} := E_t^{\mathbb{Q}^T}(R_{t,T}^2), W_{t,T} := E_t^{\mathbb{Q}^T}(R_{t,T}^3), X_{t,T} := E_t^{\mathbb{Q}^T}(R_{t,T}^4), \quad (4)$$

which are computed by the standard procedure from Carr and Madan (2001). After having defined the function $Q(K)$ as

$$Q(K) \equiv \begin{cases} P_{t,T}(K) & K \leq F_{t,T} \\ C_{t,T}(K) & \text{otherwise,} \end{cases} \quad (5)$$

where $P_{t,T}(K)$ and $C_{t,T}(K)$ are the put and call prices with strike K , we get the following expressions for the non-central implied moments

$$\mu_{t,T} = -\frac{1}{p_{t,T}} \int_0^{F_{t,T}} \frac{Q(K)}{K^2} dK \quad (6)$$

$$V_{t,T} = \frac{1}{p_{t,T}} \int_0^{F_{t,T}} \frac{2 - 2 \log\left(\frac{K}{F_{t,T}}\right)}{K^2} Q(K) dK \quad (7)$$

$$W_{t,T} = \frac{1}{p_{t,T}} \int_0^\infty \frac{\left(2 - 3 \log\left(\frac{K}{F_{t,T}}\right)\right) \log\left(\frac{K}{F_{t,T}}\right)}{K^2} Q(K) dK \quad (8)$$

$$X_{t,T} = \frac{1}{p_{t,T}} \int_0^\infty \frac{\left(3 - 4 \log\left(\frac{K}{F_{t,T}}\right)\right) \log^2\left(\frac{K}{F_{t,T}}\right)}{K^2} Q(K) dK, \quad (9)$$

and, eventually, we retrieve the central implied moments as follows:

$$\text{Var}_t^{\mathbb{Q}^T}(R_{t,T}) \equiv E_t^{\mathbb{Q}^T}((R_{t,T} - \mu_{t,T})^2) = V_{t,T} - \mu_{t,T}^2 \quad (10)$$

$$\text{Skew}_t^{\mathbb{Q}^T}(R_{t,T}) \equiv E_t^{\mathbb{Q}^T}\left(\left(\frac{R_{t,T} - \mu_{t,T}}{\sqrt{\text{Var}_t^{\mathbb{Q}^T}(R_{t,T})}}\right)^3\right) = \frac{W_{t,T} - 3V_{t,T}\mu_{t,T} + 2\mu_{t,T}^3}{\text{Var}_t^{\mathbb{Q}^T}(R_{t,T})^{3/2}} \quad (11)$$

$$\text{Kurt}_t^{\mathbb{Q}^T}(R_{t,T}) \equiv E_t^{\mathbb{Q}^T}\left(\left(\frac{R_{t,T} - \mu_{t,T}}{\sqrt{\text{Var}_t^{\mathbb{Q}^T}(R_{t,T})}}\right)^4\right) = \frac{X_{t,T} - 4W_{t,T}\mu_{t,T} + 6V_{t,T}\mu_{t,T}^2 - 3\mu_{t,T}^4}{\text{Var}_t^{\mathbb{Q}^T}(R_{t,T})^2}. \quad (12)$$

Then, the derivation of the realized analogues for the central moments is based on a buy and hold strategy on the non-central ones, where a g twice-differentiable function on $F_{t,T}$ is traded. In our specific case, we are interested in the realized measures for

$VAR_{t,T}^{Q_T}$, $SKEW_{t,T}^{Q_T}$, and $KURT_{t,T}^{Q_T}$, thus we pick the following choices for the function g :

$$g^\mu(x) \equiv \log\left(\frac{x}{F_{t,T}}\right), g^V(x) \equiv \log\left(\frac{x}{F_{t,T}}\right)^2, g^W(x) \equiv \log\left(\frac{x}{F_{t,T}}\right)^3, g^X(x) \equiv \log\left(\frac{x}{F_{t,T}}\right)^4, \quad (13)$$

along with the floating leg:

$$RV_{t,T}^\xi = g^\xi(F_{t,T}). \quad (14)$$

for $\xi \in \{\mu, V, W, X\}$. Trading $R_{t,T}$ in Eqs. (10) to (12), we finally get

$$RV_{t,T}^{Var}(N) = RV_{t,T}^V(N) - 2RV_{t,T}^\mu(N)\mu_{t,T} + \mu_{t,T}^2, \quad (15)$$

$$RV_{t,T}^{Skew}(N) = \frac{RV_{t,T}^W(N) - 3RV_{t,T}^V(N)\mu_{t,T} + 3RV_{t,T}^\mu(N)\mu_{t,T}^2 - \mu_{t,T}^3}{\text{Var}_t^{Q_T}(R_{t,T})^{3/2}}, \quad (16)$$

$$RV_{t,T}^{Kurt}(N) = \frac{RV_{t,T}^X(N) - 4RV_{t,T}^W(N)\mu_{t,T} + 6RV_{t,T}^V(N)\mu_{t,T}^2 - 4RV_{t,T}^\mu(N)\mu_{t,T}^3 + \mu_{t,T}^4}{\text{Var}_t^{Q_T}(R_{t,T})^2}. \quad (17)$$

Accordingly, all we need in order to trade these quantities are just forward and options written on $F_{t,T}$.

We argued so far that designing a correct measure for realized central moments is a critical issue for the exact definition of risk premia. This matter is particularly important in the case of skewness swaps. The reason is straightforward: unlike variance indexes and variance swap payoffs which are mostly based on a non-central moment formulas, the only implied skewness index developed in the financial industry, the CBOE SKEW, is constructed as a central moment measure. Thus, as the forward-price of realized volatility is the VIX index, the only realized skewness measure having the SKEW index as forward price is $RV_{t,T}^{Skew}(N)$.

1.1 Risk aversion and tradeable central moments

Our approach of centering both the legs of the central moments swap under the same measure may pose additional questions on the factors influencing the relation between physical and risk-neutral moments. More specifically, we are interested in verifying whether these factors are different when the P central moments are centered under the

Q_T measure. Hence, conforming to the approach of BKM, we perform a similar analysis on the relation between risk-neutral and physical skewness. Notably, they assume the existence of an investor with power utility function, which implies the pricing kernel to be represented by the exponential function $e^{-\gamma R_{t,T}}$, where γ is the coefficient of relative risk aversion. In this framework, they show that the risk-neutral and physical skewness are related according to the approximation¹

$$\text{Skew}_t^{Q_T}(R_{t,T}) \approx \text{Skew}_t^P(R_{t,T}) \left(1 + \frac{3\text{E}_t^P(R_{t,T}^3)\gamma}{2\text{E}_t^P(R_{t,T}^2)} \right) \quad (18)$$

$$- \gamma (\text{Kurt}_t^P(R_{t,T}) - 3) \sqrt{\text{Var}_t^P(R_{t,T})}, \quad (19)$$

where $\text{Var}_t^P(R_{t,T})$, $\text{Skew}_t^P(R_{t,T})$ and $\text{Kurt}_t^P(R_{t,T})$ are the return variance skewness and kurtosis under the physical measure. On the other hand, the equation connecting the measure for realized skewness introduced above and its implied equivalent is illustrated in the following Proposition:

Proposition 1.2. *Up to a first order of γ , the risk-neutral skewness of log-returns is analytically linked to its physical tradeable counterpart according to the approximation:*

$$\text{Skew}_t^{Q_T}(R_{t,T}) \approx \text{E}_t^P(RV_{t,T}^{Skew}) - \gamma \text{E}_t^P(RV_{t,T}^{Kurt}) \frac{\text{E}_t^P(RV_{t,T}^{Var})}{\text{Var}_t^{Q_T}(R_{t,T})^{3/2}}, \quad (20)$$

Proof. From the Appendix of BKM we know that:

$$\mu_{t,T} \approx \text{E}_t^P(RV_{t,T}^\mu) - \gamma \text{E}_t^P(RV_{t,T}^V) \quad (21)$$

$$V_{t,T} \approx \text{E}_t^P(RV_{t,T}^V) - \gamma \text{E}_t^P(RV_{t,T}^W) \quad (22)$$

$$W_{t,T} \approx \text{E}_t^P(RV_{t,T}^W) - \gamma \text{E}_t^P(RV_{t,T}^X), \quad (23)$$

and, assuming that $\text{E}_t^P(RV_{t,T}^\mu) = 0$, then:

$$\text{Var}_t^{Q_T}(R_{t,T}) = V_{t,T} - \mu_{t,T}^2 \approx \text{E}_t^P(RV_{t,T}^V) - \gamma \text{E}_t^P(RV_{t,T}^W) + \gamma^2 \text{E}_t^P(RV_{t,T}^V)^2. \quad (24)$$

¹The correct approximation is slightly different from equation (13) in Bakshi et al. (2003). We demonstrate that in Appendix A.

By exploiting Equation (24) we get:

$$\begin{aligned}
E_t^P(RV_{t,T}^{Skew}) &= \frac{E_t^P(RV_{t,T}^W) - 3E_t^P(RV_{t,T}^V)\mu_{t,T} + 3E_t^P(RV_{t,T}^\mu)\mu_{t,T}^2 - \mu_{t,T}^3}{\text{Var}_t^{Q_T}(R_{t,T})^{3/2}} \\
&= \frac{W_{t,T} + \gamma E_t^P(RV_{t,T}^X) - 3(V_{t,T} + \gamma E_t^P(RV_{t,T}^W))\mu_{t,T} - \mu_{t,T}^3}{\text{Var}_t^{Q_T}(R_{t,T})^{3/2}} + o(\gamma) \quad (25) \\
&\approx \text{Skew}_t^{Q_T}(R_{t,T}) + \gamma E_t^P(RV_{t,T}^{Kurt}) \frac{E_t^P(RV_{t,T}^{Var})}{\text{Var}_t^{Q_T}(R_{t,T})^{3/2}}.
\end{aligned}$$

□

A comparison between the expression in Proposition 1.2 and the corresponding approximation in BKM reveals that considering a tradeable measure for realized skewness does not modify the three sources determining negative risk-neutral skewness: negative skew in the physical distribution, risk aversion and kurtosis. However, the latter has an impact in absolute terms and not in the form of excess kurtosis as in the case of BKM. This finding has significant impact for the relative risk aversion evaluated empirically in Section 3. The coefficient multiplying risk aversion and kurtosis, given by $E_t^P(RV_{t,T}^{Var})/\text{Var}_t^{Q_T}(R_{t,T})^{3/2}$, has also a slightly different structure, as a result of centring under the Q_T measure.

The approximations in Equation (18) and in Equation (24) can potentially have an interesting usage in terms of risk aversion estimation. When conditional moments are evaluated, the risk aversion of an investor with power utility preferences might be gauged by simply solving for γ . Coherently, we define two different measures for this parameter, γ_{untrad} and γ_{trad} , where the first one is based on the standard approach of centring moments with respect to $E_t^P(R_{t,T})$, while the second one results from our

tradeable scheme:

$$\gamma_{untrad} \approx - \frac{E_t^P (R_{t,T}^2) \left(\text{Skew}_t^P (R_{t,T}) - \text{Skew}_t^{Q_T} (R_{t,T}) \right)}{\left(3E_t^P (R_{t,T}^3) \text{Skew}_t^P (R_{t,T}) - 2E_t^P (R_{t,T}^2) \left(\text{Kurt}_t^P (R_{t,T}) - 3 \right) \sqrt{\text{Var}_t^P (R_{t,T})} \right)}, \quad (26)$$

$$\gamma_{trad} \approx \frac{\left(E_t^P (RV_{t,T}^{Skew}) - \text{Skew}_t^{Q_T} (R_{t,T}) \right) \text{Var}_t^{Q_T} (R_{t,T})^{3/2}}{E_t^P (RV_{t,T}^{Kurt}) E_t^P (RV_{t,T}^{Var})}. \quad (27)$$

Unfortunately, both the proxies for γ are tightly linked to a precise estimation of the conditional central moments under P , an arduous task, in particular when measuring kurtosis (a detailed discussion on the effectiveness of central moments to measure skewness and kurtosis is contained in [Kim and White \(2004\)](#)).

Therefore, we devise an alternative methodology to circumvent this problem and identify a different proxy for γ , possibly involving only the expected P -conditional variance, which can be predicted to a considerable extent by its implied counterpart, as shown by [Kozhan et al. \(2013\)](#), and the implied moments recovered from options. The result is summarized by the following proposition:

Proposition 1.3. *An alternative γ_{trad} approximation is given by:*

$$\gamma_{trad} \approx \frac{E_t^P (RV^{Var}) - V_{t,T} - \mu_{t,T}^2}{W_{t,T}}. \quad (28)$$

Proof. We follow again the approach of BKM but in this case we aim at exponentially tilting the forward measure Q_T . In particular, in a power utility economy we have the forward and physical density of the (t, T) period return, denoted by $q(R_{t,T})$ and $p(R_{t,T})$ respectively, which are related by the Radon-Nikodym theorem as follows:

$$p(R_{t,T}) = \frac{e^{\gamma R_{t,T}} q(R_{t,T})}{\int e^{\gamma R_{t,T}} q(R_{t,T}) dR_{t,T}}. \quad (29)$$

Then, we suppress the dependence on (t, T) for convenience and define the moment

generating function $M_q(\lambda)$ of $q(R)$, for any real number γ , by:

$$M_q(\lambda) = \int_{-\infty}^{\infty} e^{\lambda R} q(R) dR = 1 + \lambda\mu + \frac{\lambda^2}{2}V + \frac{\lambda^3}{6}W + \frac{\lambda^4}{24}X + o(\lambda^4). \quad (30)$$

From Equation (29) we get the following expression for the moment generating function $M_p(\lambda)$ of $p(R)$:

$$M_p(\lambda) = \int_{-\infty}^{\infty} e^{\lambda R} p(R) dR = \frac{\int_{-\infty}^{\infty} e^{(\lambda+\gamma)R} q(R) dR}{\int_{-\infty}^{\infty} e^{\gamma R} q(R) dR} = \frac{M_q(\lambda + \gamma)}{M_q(\gamma)}. \quad (31)$$

Now, without loss of generality, we can assume that $p[R]$ has been mean shifted ($E_t^P(RV_{t,T}^\mu) = 0$). After computing the first and second derivatives with respect to λ for both the left and right hand side of Equation (31) and imposing $\lambda = 0$, up to a first order effect of γ , we end up with the following relationships:

$$E_t^P(RV^\mu) = 0 \approx \frac{\mu + \gamma V}{1 + \gamma\mu} \rightarrow \mu \approx -\gamma V, \quad (32)$$

$$E_t^P(RV^V) \approx \frac{V + \gamma W}{1 + \gamma\mu} + \mu^2 = \frac{V + \gamma W}{1 + \gamma^2\mu V} \approx V + \gamma W. \quad (33)$$

Eventually, we are able to calculate an approximation for the expected realized variance:

$$E_t^P(RV^{Var}) = E_t^P(RV^V) - 2E_t^P(RV^\mu)\mu + \mu^2 \approx V + \gamma W + \mu^2, \quad (34)$$

and, solving for γ , the proof is completed. \square

The approximation described in Equation (28) together with the untradeable equivalent in Equation (26) will be tested on data regarding forward and options on the S&P 500 in Section 3.

In summary, this section presents a way to trade central moments appropriately, by providing a strict definition of risk premium which holds for every strategy (not only for central moments) and is associated to the concept of tradeability. In the next section we prove that, in the case of central moments, sample moments should not be adopted to compute risk premia even under the unrealistic assumption of an investor able to build a strategy to trade them.

1.2 Non-central Moments and Predictions

Since we and Bakshi have assumed that first moment under $p=0$ we develop this section. An interesting application stems from the fact that forward-neutral moments are fully conditional and therefore lend themselves to be used in predictive regressions. New literature evolves around this. For example, turning around Eq. (52) we have (denoting by $\kappa_n := \mathbb{E}_t^{Q^T} [R]$ and $\bar{\kappa}_n := \mathbb{E}_t^P [R]$)²

$$\bar{\kappa}_1 = \kappa_1 + \gamma (\kappa_2 - \kappa_1^2) + o(\gamma^2) \quad (37)$$

$$= \gamma Var^{Q^T} - 2VIX^2 + o(\gamma). \quad (38)$$

The difference between VIX^2 and Var^{Q^T} is related to forward-neutral skewness. Both, log returns, as well as the skewness of log returns are directional moments. An interesting application of this would be to look for γ which yields a zero-expected log return under P . Up to terms of order $O(\gamma^2)$ this would mean

$$\gamma = \frac{2VIX^2}{Var^{Q^T}} \quad (39)$$

To first order this would also imply a zero equity premium (EP). For the risk premium on VIX^2 this would mean

$$VRP = -\bar{\kappa}_1 + EP + \kappa_1 \approx -\bar{\kappa}_1 + \kappa_1 = \gamma (\kappa_1^2 - \kappa_2) + O(\gamma^2) = -\gamma Var^{Q^T} + O(\gamma^2) \quad (40)$$

so that the variance premium itself is proportional to forward-neutral variance and negative. To third order

$$VRP = \gamma (\kappa_1^2 - \kappa_2) + \frac{1}{2}\gamma^2 (-2\kappa_1^3 + 3\kappa_2\kappa_1 - \kappa_3) + O(\gamma^3) \quad (41)$$

$$= -\gamma Var^{Q^T} - \frac{1}{2}\gamma^2 (Var^{Q^T})^{3/2} Skew^{Q^T} + O(\gamma^3) \quad (42)$$

² Taking this further one order we would have

$$\bar{\kappa}_1 = \kappa_1 + \gamma (\kappa_2 - \kappa_1^2) + \frac{1}{2}\gamma^2 (2\kappa_1^3 - 3\kappa_2\kappa_1 + \kappa_3) + O(\gamma^3) \quad (35)$$

$$= \gamma Var^{Q^T} - 2VIX^2 + \frac{1}{2}\gamma^2 (Var^{Q^T})^{3/2} Skew^{Q^T} + O(\gamma^3). \quad (36)$$

2 Sample central moments and risk premia

A broad number of recent findings on variance and skewness measurement has cast doubt on the ability of sample moments to estimate precisely the actual return variation and asymmetry (see, for example, [Andersen et al. \(2001\)](#), [Barndorff-Nielsen \(2002\)](#) and [Neuberger \(2012\)](#)). Nonetheless, there is an additional issue intimately related to central moments which justifies our idea of adopting realized tradeable measures for variance, skewness and kurtosis instead of their sample counterparts. To frame the problem, we can think about a claim that pays $\mu_{t,T}^P$, the t -conditional first moment of X_T has the conditional density of X as underlying. This conditional density is not directly observable, and there exists no technology which could attain it in absence of a model. In practice, it is common to make inferences about the P distribution by looking at returns over many periods. The question then arises about the relation between the P moments estimated from a long series of returns, and the average conditional Q_T moments computed over the same period from option prices. The following propositions show that this common practice is valid when considering non-central moments but is definitely not well-grounded if central moments are the object of the analysis.

Proposition 2.1. *Significant differences between the n th average implied non-central moment and its population counterpart imply the presence of n th moment risk premium.*

Proof. In an economy with no risk premia ($P = Q_T$) investors will not get any excess return for trading non-central moments:

$$E_t^P (R_{t,T}^n) - E_t^{Q_T} (R_{t,T}^n) = 0. \quad (43)$$

If we calculate the unconditional expectation on both sides we obtain:

$$E^P (E_t^P (R_{t,T}^n)) - E^P (E_t^{Q_T} (R_{t,T}^n)) = 0, \quad (44)$$

and by applying the law of iterated expectation for the first term on the left-hand side and rearranging:

$$E^P (R_{t,T}^n) = E^P (E_t^{Q_T} (R_{t,T}^n)), \quad (45)$$

so that mean option-implied moments converge to the (unconditional) population mo-

ment. □

However, the same conclusion does not apply to the common definition of variance, skewness and kurtosis since the equality does not hold.

Proposition 2.2. *Significant differences between average implied variance/skewness/kurtosis and its population counterpart does not demonstrate the presence of variance/skewness/kurtosis risk premium.*

Proof. Here we provide the proof for the variance, whereas skewness and kurtosis are treated in Appendix B. In an economy with no risk premia ($P = Q_T$) first and second moments under the two measures will be equal:

$$\text{Var}_t^P (R_{t,T}) - \text{Var}_t^{Q_T} (R_{t,T}) = 0. \quad (46)$$

If we calculate the unconditional expectation on both sides we obtain:

$$\mathbb{E}^P \left(\text{Var}_t^P (R_{t,T}) \right) - \mathbb{E}^P \left(\text{Var}_t^{Q_T} (R_{t,T}) \right) = 0, \quad (47)$$

and, by applying the law of total variance for the first term on the left-hand side, and rearranging:

$$\text{Var}^P (R_{t,T}) = \mathbb{E}^P \left(\text{Var}_t^{Q_T} (R_{t,T}) \right) + \text{Var}^P \left(\mathbb{E}_t^P (R_{t,T}) \right). \quad (48)$$

□

Indeed, it is easy to think of cases where there are no risk premia and population moments differ from average implied moments, as shown in the following example.

Example 2.1. Take the discrete-time model

$$R_{t,T} = -\frac{1}{2}v_t + \sqrt{v_t}\varepsilon_{t,T}, \quad (49)$$

where v_t is known at the beginning of period t and $\varepsilon_{t,T}$ is standard normal. In the absence of risk premia, the implied variance in period t is v_t , so the mean implied variance is \bar{v} (the unconditional mean variance). The population variance is

$$\bar{v} + \mathbb{V}[v]/4. \quad (50)$$

3 Empirical analysis

3.1 Data and methodology

The empirical analysis concentrates on the S&P 500 central moment risk premiums. Data on this index are taken from OptionMetrics and cover a period from March 1996 to January 2015. We construct three trading strategies, for a monthly, 3-months and 6-months time grid respectively. Essentially, for every date in the sample, we collect the data for the three mentioned maturities starting from March instead of January in order to be able to pick the three and six-month maturity, being the S&P500 options always traded for the March, June, September and December third Friday expiration. The variables in the dataset consist of bid and ask quotes, forward price for the corresponding expiry, interest rate, strike price, implied volatility and volume. Option prices are computed as midpoints of the last bid-ask quote and all entries with non standard-settlements, with implied volatility between 0.001 and 9 and with no transactions occurred in the given date are filtered out. The standard no-arbitrage filters for European options are also applied.

Since the purpose of this study is to obtain all completely tradeable quantities, we do not implement any type of interpolation, neither for computing returns for time horizons which are not in line with the options expiration structure, nor for retrieving the option prices above/below the highest/lowest traded strike (see [Carr and Wu \(2009\)](#) for one example of this procedure). In this respect, the number of strikes available in the market represents a statistic particularly worthy of attention. Indeed, the simulation experiment in [Jiang and Tian \(2005\)](#) suggests that truncation and discretization error of implied volatility estimation, arising from the fact that a continuum of options is not actually traded, are negligible if options for at least 20 different strikes are traded. Thus, we investigate whether the sample resulting from the criteria shown previously satisfies this condition by reporting in [Figure 1](#) the number of strikes for which at least one contract was negotiated at every date. For the one-month expirations only a few observations stand below the twenty-strikes threshold while the errors might be more pronounced for the three and six-month time horizon. Therefore, from the point of view of estimation accuracy, the implied variance, skewness and kurtosis retrieved in this section should be treated with caution. However, we remind that the top priority

of this paper does not consist in computing with the maximum possible precision these moments but in designing the best feasible strategies to trade them.

3.2 Central moment swaps on the S&P500

Prices of forward and options are used to compute realized central moments, according to the technology introduced in Section 1, jointly with the corresponding BKM implied moments and with the rolling sample central moments. In order to evaluate quantities comparable with sample moments, both the realized and the implied legs are calculated after a burn-in period of three, five and eight years, for the monthly, three-month and six-month maturity. The discrepancy among the lengths of the burn-in periods is due to the different number of observations within a year for the three maturities. Shrinking or enlarging the rolling window for the sample moments calculation does not substantially affect the results.

Summary statistics on the three central moment estimates are reported in Table 1. The average level of realized variance is lower than its implied counterpart for all the maturities, entailing a negative variance risk premium, consistently with several previous findings in the literature (e.g, Bakshi and Kapadia (2003), Carr and Wu (2009) and Bondarenko (2014)). On the contrary, both the average realized and sample skewness are higher (closer to zero) than average implied skewness, in line both with Equation (24) and with the skew swap risk premiums of Neuberger (2012) and Kozhan et al. (2013). Unsurprisingly, most of the realized moments are less volatile than their implied equivalents, being these Q_T expectations, but both show a higher variability than sample moments, since the latter are estimates depending on a rolling window.

Evolution of central moments along time is displayed in Figures 2 to 4. The charts clearly confirm the idea that realized central moments are more appropriate than sample moments for measuring risk premia. Realized variance, for instance, has very similar fluctuations to the ones of the BKM because both do not depend on past observations. This fact is particularly evident during crisis events since the more visible peaks in the paths almost coincide. In November 2009, when Lehmann Brothers went bankrupt, the sample variance was not affected as drastically as the realized and implied counterparts by the abnormal returns observed in that period. Skewness and kurtosis are also

	Sample	Implied	Realized
<i>Monthly returns</i>			
Variance (x 100)	0.285 (0.178)	0.445 (0.599)	0.282 (0.671)
Skewness	-0.783 (0.503)	-2.186 (0.865)	-0.524 (3.016)
Kurtosis	4.001 (1.035)	17.631 (18.248)	1.881 (7.425)
<i>Three monthly returns</i>			
Variance (x 100)	0.850 (0.350)	1.404 (1.113)	0.819 (1.501)
Skewness	-0.656 (0.482)	-2.339 (2.033)	-0.200 (1.608)
Kurtosis	3.297 (1.173)	22.348 (43.753)	0.914 (2.894)
<i>Six monthly returns</i>			
Variance (x 100)	2.541 (0.400)	2.481 (1.943)	2.131 (4.510)
Skewness	-0.939 (0.202)	-2.141 (0.533)	-0.458 (3.215)
Kurtosis	3.482 (1.0894)	10.888 (5.048)	1.943 (7.559)

Table 1: central moments statistics. The table reports mean and standard deviation (between parenthesis) for the implied, the realized and rolling sample central moments of the S &P500. Both option and S&P 500 data is available from March 1996 up to January 2015. Calculation of the three statistics starts in March 1999.

following a similar pattern. In particular, implied kurtosis lies above sample kurtosis also during market shocks, confirming a rather small correlation between the two measures. On the other hand, realized kurtosis tends to spike in market turmoil so that the kurtosis risk premium calculated with our measure is in agreement with the intrinsic nature of a kurtosis contract as an insurance claim covering huge (and mostly negative) returns. Nevertheless, we remind that estimating kurtosis represents a demanding task. For instance, values of implied kurtosis are likely to be misleading since the truncation error has a more severe impact on fourth moment, being this one directly related to tail risk, which is only partially observable from a bounded set of traded options.

3.3 Estimating risk aversion from central moment swaps

As documented by [Jackwerth \(2000\)](#), the risk neutral density of the S&P500 estimated from option prices is more left-skewed than the physical distribution measured with historical data, suggesting a positive empirical skewness risk premium, consistently also with the results exhibited in Table 1. As argued in Section 1, BKM identify, in a power utility economy, a relation between the skewness risk premium and the expected excess kurtosis, whereas, we derive an analogue equation for tradable central moments depending only on the absolute level of the expected kurtosis. The difference is subtle but the impact on the coefficient of relative risk aversion γ is not trivial. Indeed, market conditions in which the skewness risk premium is positive even when the expected excess kurtosis is negative cannot be excluded a priori. For example, Figures 2 to 4 show that in the period of time starting from the end of the dot-com crisis and ending with the subprime mortgage crisis (2002-2007) both the realized and sample kurtosis were often lower than three, but the skewness risk premium was still positive. This stylized fact suggests an experiment on the evaluation of risk aversion using both the proxies γ_{trad} and γ_{untrad} presented in Section 1. Specifically, we retrieve values for the relative risk aversion given by Equations (26) and (28), using demeaned returns. To accomplish this purpose, we infer the conditional expectation $E_t^P (RV_{t,T}^{Var})$ through the linear model:

$$RV_{t,T}^{Var} = \alpha + \beta \text{Var}_t^{Q_T} (R_{t,T}) + \epsilon_t \quad (51)$$

for all the maturities considered. As shown by [Kozhan et al. \(2013\)](#), the predictable power of implied second moments on their realized equivalent seems to be considerable, providing sufficiently solid grounds for the selection of this model. The OLS estimator regression is run in a predictive framework where the first estimate is obtained after the burn-in periods already described, and the subsequent ones are computed by expanding the window at every date with one additional observation. P-values for the slope level β together with the R^2 are shown in Figures 5 and 6. In order to achieve the maximum degree of accuracy for the assessment of $E_t^P (RV_{t,T}^{Var})$ we deem as reliable regressions only the ones with p-values lower than 1%. Therefore, we get rid of all the observations before December 2006 and December 2011 for the one-month and three-months maturity respectively, and of the complete sample for what concerns the

six-month maturity. After having applied this filter, we can infer $E_t^P(RV_{t,T}^{Var})$ from the estimated coefficients α and β and from the explanatory variable $\text{Var}_t^{Q^T}(R_{t,T})$ (Figures 7a and 7b). Ultimately, we are able to obtain our proxy γ_{trad} , reported in Figures 7e and 7f. This parameter is included in a range between 0.192 and 5.953 for the monthly maturity, and between 1.196 and 2.318 for the 3-month maturity, never violating the positivity constraint. These values are consistent with the upper bound $\gamma \leq 10$ suggested by Mehra and Prescott (1985). The zero line is never crossed because, as already illustrated in this paper, the literature has broadly demonstrated that both the expected variance risk premium and the third moment of the risk-neutral distribution tend to be negative. The expression in Equation (28) is basically a ratio between these two quantities.

The estimation of γ_{untrad} cannot follow the same scheme due to the difficulties in forecasting kurtosis. Hence, to measure $\text{Skew}_t^P(R_{t,T})$ and $\text{Kurt}_t^P(R_{t,T})$ we use sample moments. Results shown in Figures 7c and 7d are patently in contradiction with the assumption $\gamma > 0$, presenting further evidence that sample moments are utterly unsuitable for central moment risk premia assessment.

4 Conclusions

In this paper we present for the first time in the literature, to the best of our knowledge, a tradeable strategy for computing risk premia depending on central moments. Besides being replicable with traded instruments, our strategy creates an exact correspondence between floating and fixed legs, since realized moments reflect BKM prices perfectly. This aspect is crucial, because subtracting some realized quantity from some other unrelated implied measure might undoubtedly not be sufficient to classify the difference as a risk premium.

After defining the realized measure to construct central moment swaps, we also derive a formula to identify the coefficient of relative risk aversion for a power utility investor trading these contracts. An empirical study on the S&P 500 index reveals that this formula is consistent with two of the main findings of the return distribution literature: the negative variance risk premium and the negative implied third moment. These two stylized facts explain why, unlike an analogous measure based on sample mo-

ments, our proxy for the coefficient of relative risk aversion never violates the positivity constraint of this parameter.

Furthermore, apart from tradeability issues, we demonstrate that the difference between sample central moments and their implied equivalent does not serve as a valid proxy for measuring variance, skewness and kurtosis risk premia. This theoretical fact is corroborated by the study on the S&P 500 index where implied moments fluctuations appear completely unrelated with the much less volatile trajectories of sample moments.

References

- Aït-Sahalia, Y. and A. W. Lo (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. *The Journal of Finance* 53(2), 499–547.
- Aït-Sahalia, Y. and A. W. Lo (2000). Nonparametric risk management and implied risk aversion. *Journal of econometrics* 94(1), 9–51.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens (2001). The distribution of realized stock return volatility. *Journal of financial economics* 61(1), 43–76.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of realized exchange rate volatility. *Journal of the American statistical association* 96(453), 42–55.
- Bakshi, G. and N. Kapadia (2003). Delta-hedged gains and the negative market volatility risk premium. *Review of Financial Studies* 16(2), 527–566.
- Bakshi, G., N. Kapadia, and D. Madan (2003). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies* 16(1), 101–143.
- Bakshi, G. and G. Panayotov (2010). First-passage probability, jump models, and intra-horizon risk. *Journal of Financial Economics* 95(1), 20–40.
- Barndorff-Nielsen, O. E. (2002). Econometric analysis of realized volatility and its use

- in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64(2), 253–280.
- Bondarenko, O. (2014). Variance trading and market price of variance risk. *Journal of Econometrics* 180(1), 81–97.
- Breeden, D. T. and R. H. Litzenberger (1978). Prices of state-contingent claims implicit in option prices. *Journal of business*, 621–651.
- Brillinger, D. R. (1969). The calculation of cumulants via conditioning. *Annals of the Institute of Statistical Mathematics* 21(1), 215–218.
- Britten-Jones, M. and A. Neuberger (2000). Option prices, implied price processes, and stochastic volatility. *The Journal of Finance* 55(2), 839–866.
- Carr, P. and D. Madan (2001). Towards a theory of volatility trading. *Option Pricing, Interest Rates and Risk Management, Handbooks in Mathematical Finance*, 458–476.
- Carr, P. and L. Wu (2009). Variance risk premiums. *Review of Financial Studies* 22(3), 1311–1341.
- Jackwerth, J. C. (2000). Recovering risk aversion from option prices and realized returns. *Review of Financial Studies* 13(2), 433–451.
- Jiang, G. J. and Y. S. Tian (2005). The model-free implied volatility and its information content. *Review of Financial Studies* 18(4), 1305–1342.
- Kim, T.-H. and H. White (2004). On more robust estimation of skewness and kurtosis. *Finance Research Letters* 1(1), 56–73.
- Kozhan, R., A. Neuberger, and P. Schneider (2013). The skew risk premium in the equity index market. *Review of Financial Studies* 26(9), 2174–2203.
- Mehra, R. and E. C. Prescott (1985). The equity premium: A puzzle. *Journal of monetary Economics* 15(2), 145–161.
- Neuberger, A. (2012). Realized skewness. *Review of Financial Studies* 25(11), 3423–3455.

A Proof of Equation (18)

Following the proof of Theorem 2 in the Appendix of [Bakshi et al. \(2003\)](#), we recompute the correct Taylor expansion of $\text{Skew}_t^P(R_{t,T})$. For the sake of clarity, we use here the same terminology.

Proof. We start with the following relation between P and Q moments:

$$\kappa_1 \approx \bar{\kappa}_1 - \gamma \bar{\kappa}_2, \quad \kappa_2 \approx \bar{\kappa}_2 - \gamma \bar{\kappa}_3, \quad \kappa_3 \approx \bar{\kappa}_3 - \gamma \bar{\kappa}_4, \quad (52)$$

where $\kappa_n = E_t^Q(R_{t,T}^n)$ and $\bar{\kappa}_n = E_t^P(R_{t,T}^n)$. Then:

$$\text{Skew}_t^{Q^T}(R_{t,T}) := \frac{\kappa_3 - 3\kappa_1\kappa_2 + 2\kappa_1^3}{(\kappa_2 - \kappa_1^2)^{3/2}} = \frac{\bar{\kappa}_3 - \gamma(\bar{\kappa}_4 - 3\bar{\kappa}_2^2) - 3\gamma^2\bar{\kappa}_3\bar{\kappa}_2 - 2\gamma^3\bar{\kappa}_2^3}{(\bar{\kappa}_2 - \gamma\bar{\kappa}_3 - \bar{\kappa}_2^2\gamma^2)^{3/2}}. \quad (53)$$

A first-order Taylor expansion of $\text{Skew}_t^{Q^T}(R_{t,T}, \gamma)$ around the point $\gamma_0 = 0$ leads to:

$$\text{Skew}_t^{Q^T}(R_{t,T}, \gamma) = \text{Skew}_t^{Q^T}(R_{t,T}, \gamma_0) + \frac{\partial \left(\text{Skew}_t^{Q^T}(R_{t,T}, \gamma) \right)}{\partial \gamma} \Big|_{\gamma=\gamma_0} (\gamma) + o(\gamma), \quad (54)$$

where

$$\begin{aligned} \frac{\partial \left(\text{Skew}_t^{Q^T}(R_{t,T}, \gamma) \right)}{\partial \gamma} \Big|_{\gamma=\gamma_0} &= \frac{(-\bar{\kappa}_4 + 3\bar{\kappa}_2^2 - 6\gamma\bar{\kappa}_3\bar{\kappa}_2 - 6\gamma^2\bar{\kappa}_2^3)}{(\bar{\kappa}_2 - \gamma\bar{\kappa}_3 - \bar{\kappa}_2^2\gamma^2)^{3/2}} \\ &\quad - \frac{\frac{3}{2}(-\bar{\kappa}_3 - 2\gamma\bar{\kappa}_2^2)(\bar{\kappa}_3 - \gamma(\bar{\kappa}_4 - 3\bar{\kappa}_2^2) - 3\gamma^2\bar{\kappa}_3\bar{\kappa}_2 - 2\gamma^3\bar{\kappa}_2^3)}{(\bar{\kappa}_2 - \gamma\bar{\kappa}_3 - \bar{\kappa}_2^2\gamma^2)^{5/2}}. \end{aligned} \quad (55)$$

Finally, by plugging-in (55) and (53) in (54), we obtain:

$$\begin{aligned} \text{Skew}_t^{Q^T}(R_{t,T}, \gamma) &= \frac{\bar{\kappa}_3}{\bar{\kappa}_2^{3/2}} - \frac{(\bar{\kappa}_4 - 3\bar{\kappa}_2^2)\gamma}{\bar{\kappa}_2^{3/2}} + \frac{3\bar{\kappa}_3^2}{2\bar{\kappa}_2^{5/2}}\gamma + o(\gamma) \\ &\approx \text{Skew}_t^P(R_{t,T}) \left(1 + \frac{3E_t^P(R_{t,T}^3)\gamma}{2E_t^P(R_{t,T}^2)} \right) - \gamma (\text{Kurt}_t^P(R_{t,T}) - 3) \sqrt{\text{Var}_t^P(R_{t,T})}. \end{aligned} \quad (56)$$

□

B Proof of Proposition 2.2 for Skewness and Kurtosis

B.1 Skewness

Proof. In an economy with no risk premia ($P = Q_T$):

$$\text{Skew}_t^P(R_{t,T}) - \text{Skew}_t^{Q_T}(R_{t,T}) = 0. \quad (57)$$

If we calculate the unconditional expectation on both sides we obtain:

$$E^P(\text{Skew}_t^P(R_{t,T})) - E^P(\text{Skew}_t^{Q_T}(R_{t,T})) = 0. \quad (58)$$

After defining, for a generic n , the value of the unconditional n th P -cumulant as $\kappa^n(R_{t,T})$ and knowing that

$$\begin{aligned} E^P(\text{Skew}_t^P(R_{t,T})) &= E^P\left(\frac{\kappa_t^3(R_{t,T})}{\text{Var}_t^P(R_{t,T})^{3/2}}\right) \\ &= E^P(\kappa_t^3(R_{t,T})) E^P\left(\frac{1}{\text{Var}_t^P(R_{t,T})^{3/2}}\right) + \text{Cov}^P\left(\kappa_t^3(R_{t,T}), \frac{1}{\text{Var}_t^P(R_{t,T})^{3/2}}\right) \end{aligned} \quad (59)$$

we can apply the law of total cumulance, introduced by [Brillinger \(1969\)](#), to the term $E^P(\kappa_t^3(R_{t,T}))$:

$$E^P(\kappa_t^3(R_{t,T})) = \kappa^3(R_{t,T}) - \kappa^3(E_t^P(R_{t,T})) - 3\text{Cov}^P(E_t^P(R_{t,T}), \text{Var}_t^P(R_{t,T})), \quad (60)$$

and substituting $E^P(\kappa_t^3(R_{t,T}))$ in Equation (59) we get:

$$\begin{aligned} E^P(\text{Skew}_t^P(R_{t,T})) &= E^P\left(\frac{1}{\text{Var}_t^P(R_{t,T})^{3/2}}\right) \left(\kappa^3(R_{t,T}) - \kappa^3(E_t^P(R_{t,T}))\right. \\ &\quad \left. - 3\text{Cov}^P(E_t^P(R_{t,T}), \text{Var}_t^P(R_{t,T}))\right) + \text{Cov}^P\left(\kappa_t^3(R_{t,T}), \frac{1}{\text{Var}_t^P(R_{t,T})^{3/2}}\right). \end{aligned} \quad (61)$$

Finally, replacing $E^P(\text{Skew}_t^P(R_{t,T}))$ into Equation (58), dividing both sides by $\text{Var}^P(R_{t,T})^{3/2}$ and rearranging we end up with:

$$\text{Skew}^P(R_{t,T}) = \frac{E^P(\text{Skew}_t^{Q_T}(R_{t,T}))}{E^P\left(\frac{1}{\text{Var}_t^P(R_{t,T})^{3/2}}\right) \text{Var}^P(R_{t,T})^{3/2}} + A, \quad (62)$$

where:

$$A = \frac{1}{\text{Var}^P(R_{t,T})^{3/2}} \left(\kappa^3(E_t^P(R_{t,T})) + 3\text{Cov}^P(E_t^P(R_{t,T}), \text{Var}_t^P(R_{t,T})) - \frac{\text{Cov}^P(\kappa_t^3(R_{t,T}))}{E^P\left(\frac{1}{\text{Var}_t^P(R_{t,T})^{3/2}}\right)} \right). \quad (63)$$

□

B.2 Kurtosis

Proof. As before, the assumption is the absence of risk premia ($P = Q_T$):

$$\text{Kurt}_t^P(R_{t,T}) - \text{Kurt}_t^{Q_T}(R_{t,T}) = 0 \implies E^P(\text{Kurt}_t^P(R_{t,T})) - E^P(\text{Kurt}_t^{Q_T}(R_{t,T})) = 0, \quad (64)$$

where:

$$\begin{aligned} E^P(\text{Kurt}_t^P(R_{t,T})) &= E^P\left(\frac{\kappa_t^4(R_{t,T})}{\text{Var}_t^P(R_{t,T})^2}\right) + 3 \\ &= E^P(\kappa_t^4(R_{t,T})) E^P\left(\frac{1}{\text{Var}_t^P(R_{t,T})^2}\right) + \text{Cov}^P\left(\kappa_t^4(R_{t,T}), \frac{1}{\text{Var}_t^P(R_{t,T})^2}\right) + 3. \end{aligned} \quad (65)$$

Applying the law of total cumulance to $E^P(\kappa_t^4(R_{t,T}))$ we get:

$$\begin{aligned} E^P(\kappa_t^4(R_{t,T})) &= \kappa^4(R_{t,T}) - 4\text{Cov}^P(\kappa_t^3(R_{t,T}), E_t^P(R_{t,T})) - 3\text{Var}^P(\text{Var}_t^P(R_{t,T})) \\ &\quad - E^P(\kappa_t^4(R_{t,T})) - 6\kappa(\text{Var}_t^P(R_{t,T}), E_t^P(R_{t,T})), \end{aligned} \quad (66)$$

where $\kappa(\cdot)$ is the generic expression for the joint cumulant. Then, as for the skewness proof:

$$\begin{aligned} \mathbb{E}^P(\text{Kurt}_t^P(R_{t,T})) = & \left(\kappa^4(R_{t,T}) - 4\text{Cov}^P(\kappa_t^3(R_{t,T}), \mathbb{E}_t^P(R_{t,T})) - 3\text{Var}^P(\text{Var}_t^P(R_{t,T})) \right. \\ & \left. - 6\kappa(\text{Var}_t^P(R_{t,T}), \mathbb{E}_t^P(R_{t,T})) - \mathbb{E}^P(\kappa_t^4(R_{t,T})) \right) \mathbb{E}^P\left(\frac{1}{\text{Var}_t^P(R_{t,T})^2}\right) \\ & + \text{Cov}^P\left(\kappa_t^4(R_{t,T}), \frac{1}{\text{Var}_t^P(R_{t,T})^2}\right) + 3, \end{aligned} \quad (67)$$

and $\mathbb{E}_t^P(R_{t,T})$. To conclude:

$$\text{Kurt}^P(R_{t,T}) = \frac{\mathbb{E}^P(\text{Kurt}_t^{\text{QT}}(R_{t,T}))}{\mathbb{E}^P\left(\frac{1}{\text{Var}_t^P(R_{t,T})^2}\right) \text{Var}^P(R_{t,T})^2} + B, \quad (68)$$

where:

$$\begin{aligned} B = & \frac{1}{\text{Var}^P(R_{t,T})^2} \left\{ 4\text{Cov}^P(\kappa_t^3(R_{t,T}), \mathbb{E}_t^P(R_{t,T})) + 3\text{Var}^P(\text{Var}_t^P(R_{t,T})) \right. \\ & \left. + \mathbb{E}^P(\kappa_t^4(R_{t,T})) + \frac{1}{\mathbb{E}^P(\text{Var}_t^P(R_{t,T})^2)} \left[\text{Cov}^P\left(\kappa_t^4(R_{t,T}), \frac{1}{\text{Var}_t^P(R_{t,T})^2}\right) + 3 \right] \right\}. \end{aligned} \quad (69)$$

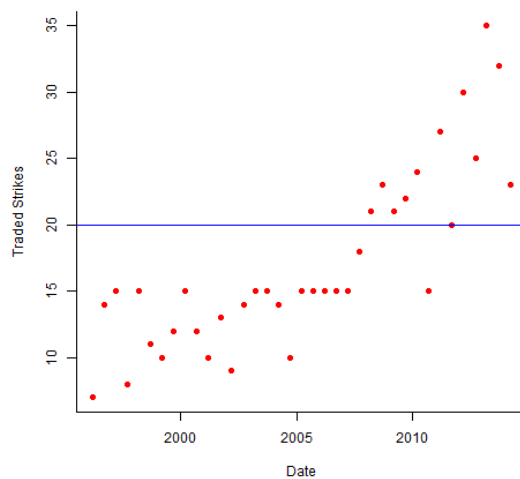
□

C Figures



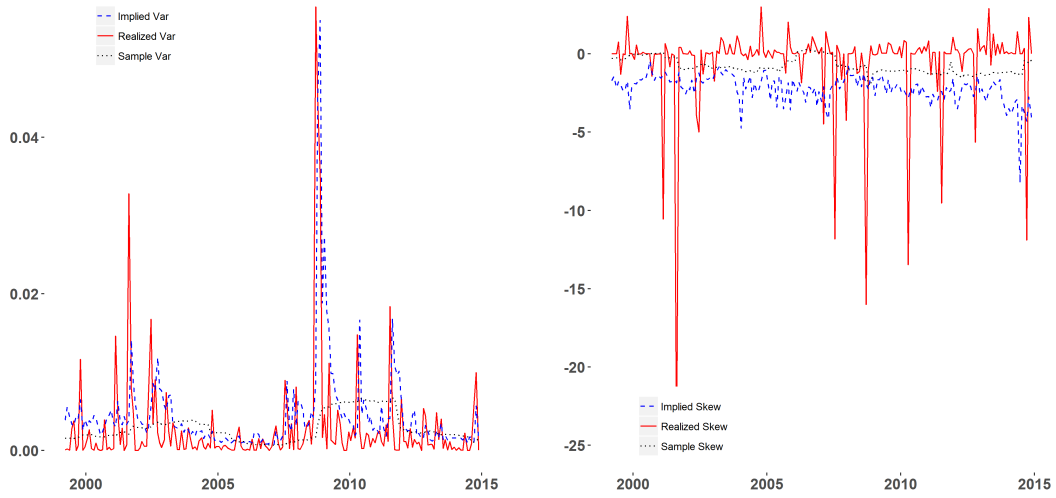
(a) Traded strikes, 1-month maturity.

(b) Traded strikes, 3-month maturity.



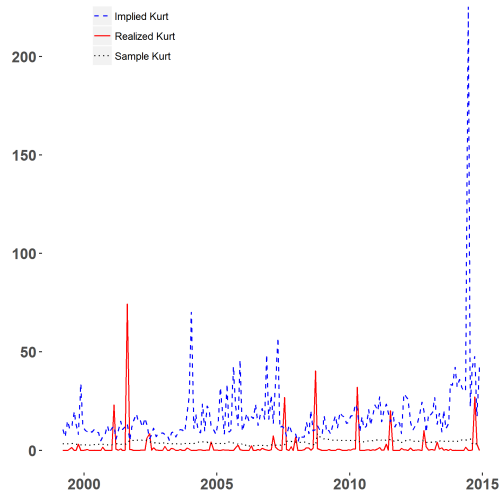
(c) Traded strikes, 6-month maturity.

Figure 1: traded strikes of the S&P 500 1996-2015. The chart shows the different strikes at which an option on the S&P 500 is traded. Maturities considered in the sample are 1-month, 3-month and 6-month. The horizontal line is the threshold indicated by [Jiang and Tian \(2005\)](#) such that, above it, the truncation and discretization errors of the VIX index are negligible.



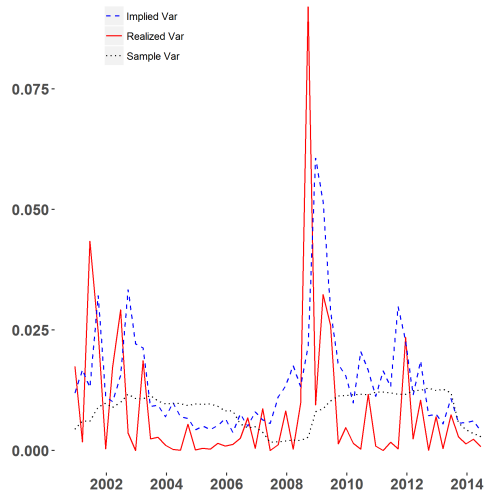
(a) Variance

(b) Skewness



(c) Kurtosis

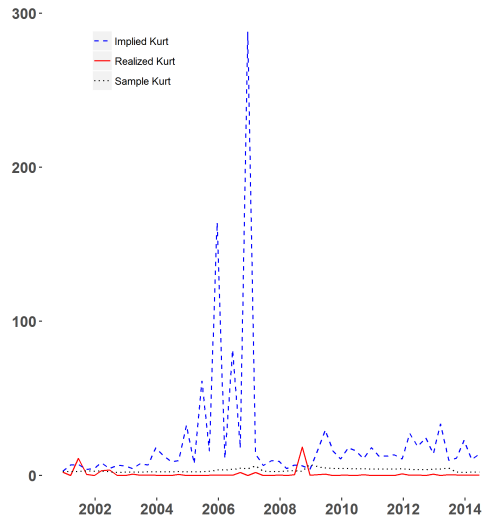
Figure 2: central moments of the S&P 500 1999-2015. The chart shows three different measures for the 3-month central moments: the implied, the realized and the 3-years rolling sample central moments of the S &P500.



(a) Variance



(b) Skewness



(c) Kurtosis

Figure 3: central moments of the S&P 500 2001-2014. The chart shows three different measures for the 6-month central moments: the implied, the realized and the 5-years rolling sample central moments of the S &P500.

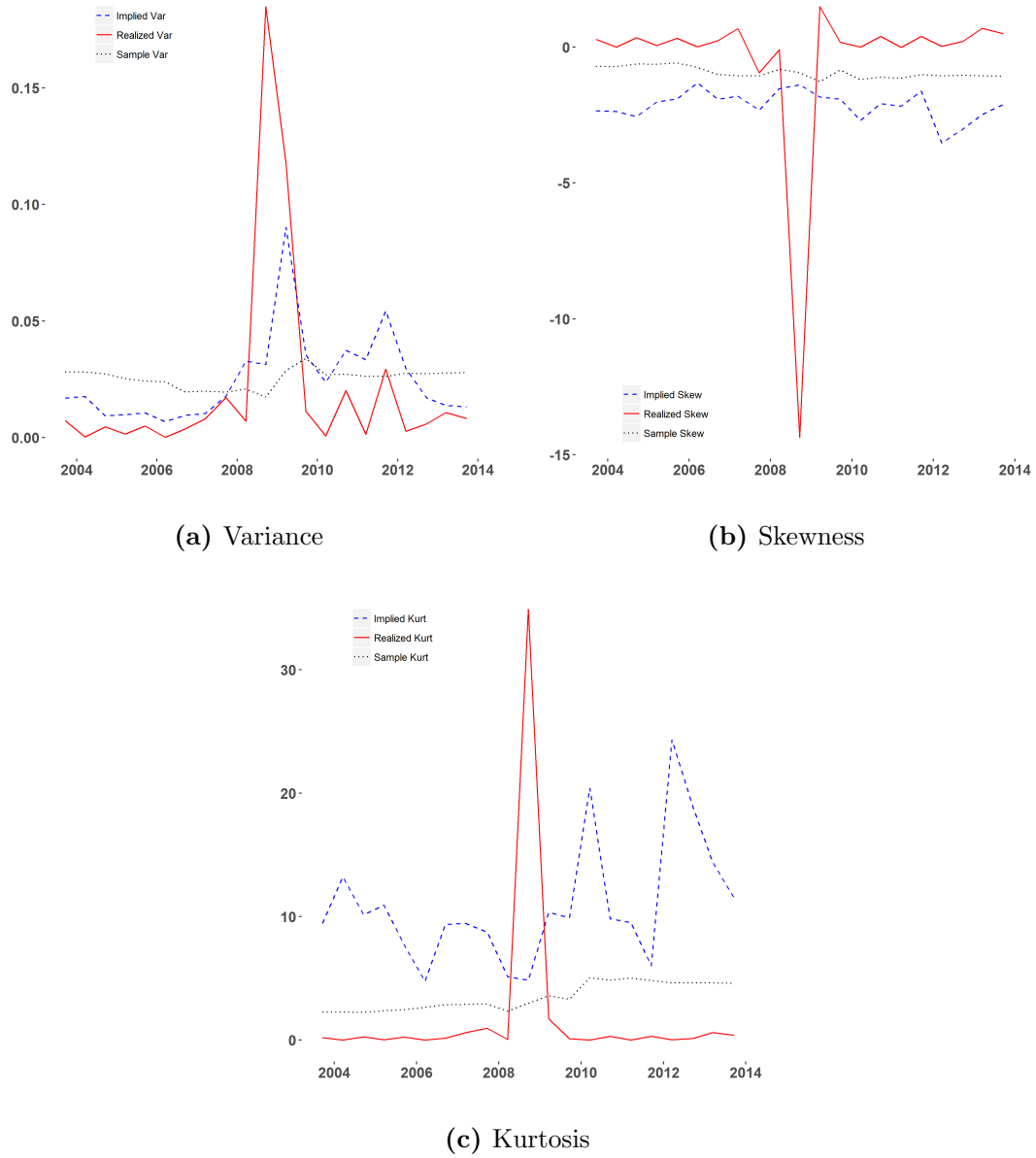
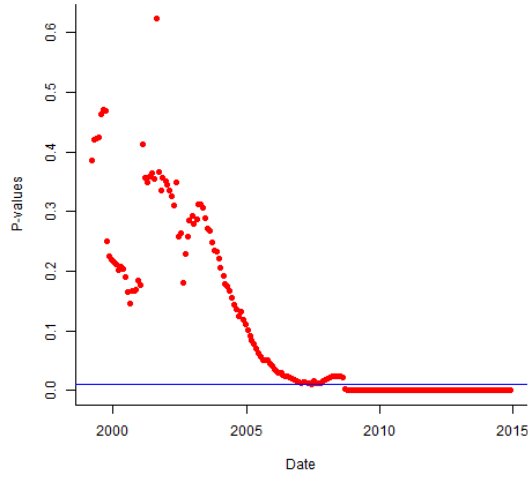
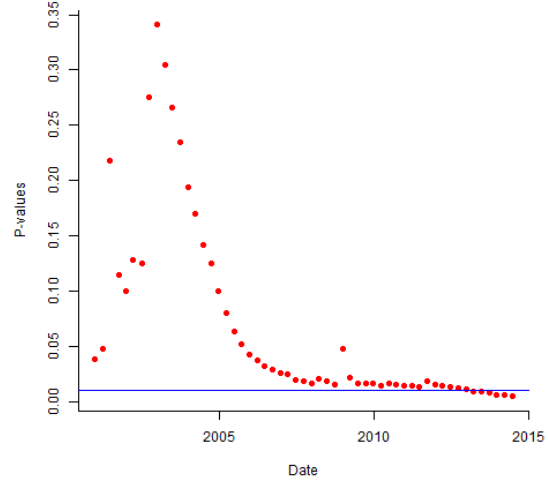


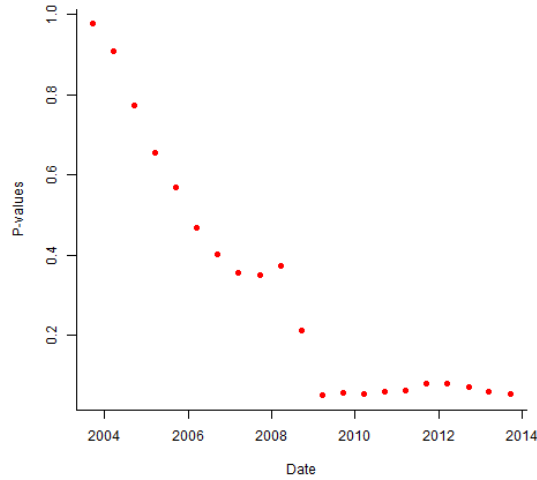
Figure 4: central moments of the S&P 500 2003-2014. The chart shows three different measures for the monthly central moments: the implied, the realized and the 8-years rolling sample central moments of the S &P500.



(a) P-values, 1-month maturity.

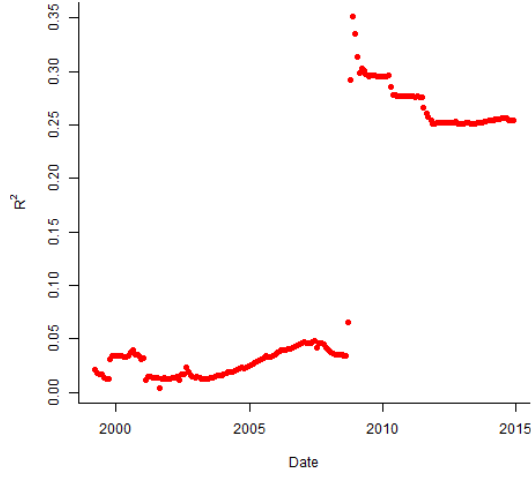


(b) P-values, 3-month maturity.

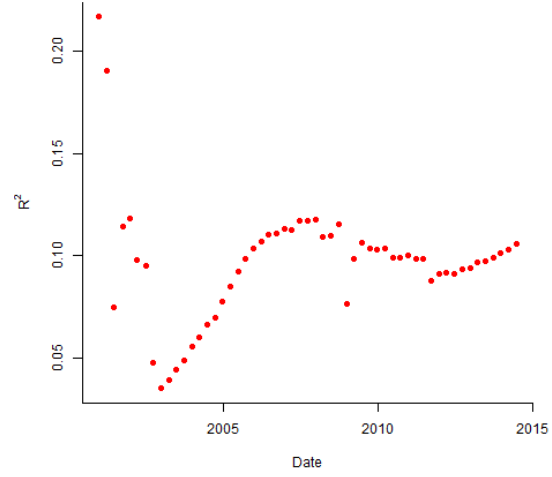


(c) P-values, 6-month maturity.

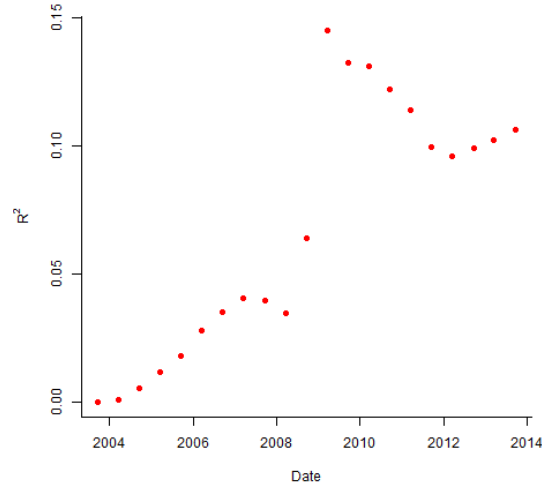
Figure 5: p-values. The chart shows the p-values of the β coefficient in the regression described by Equation (51) run for the 1-month, 3-month and 6-month realized and implied central moments of the *S&P500*. Starting from a burn-in period of 3, 5, and 8 years respectively, the standard OLS estimator is computed by adding for every date one observation (expanding window).



(a) R^2 , 1-month maturity.

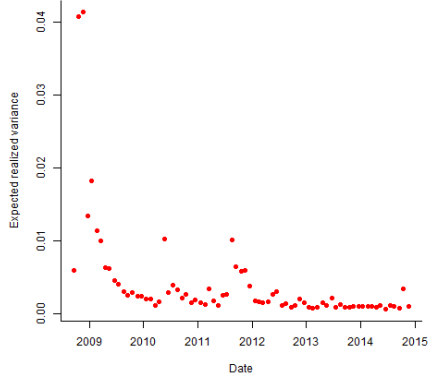


(b) R^2 , 3-month maturity.

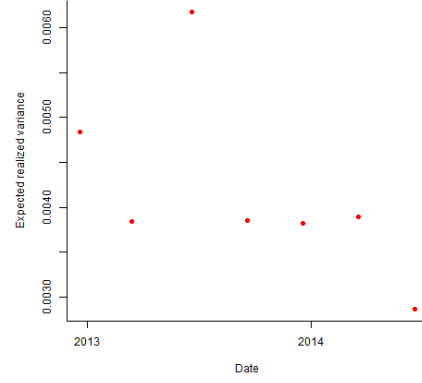


(c) R^2 , 6-month maturity.

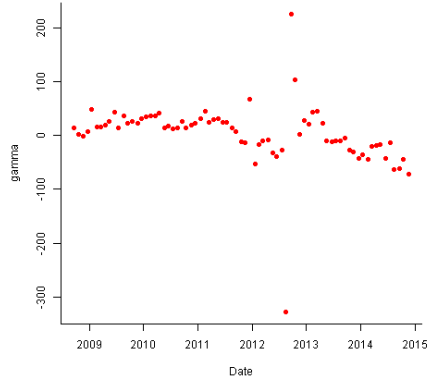
Figure 6: R^2 . The chart shows the R^2 of the regression described by Equation (51) run for the 1-month, 3-month and 6-month realized and implied central moments of the *S&P500*. Starting from a burn-in period of 3, 5, and 8 years respectively, the standard OLS estimator is computed by adding for every date one observation (expanding window).



(a) $E_t^P(RV^{Var})$, 1-month maturity



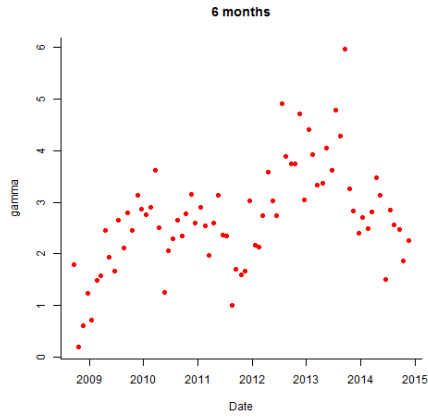
(b) $E_t^P(RV^{Var})$, 3-month maturity.



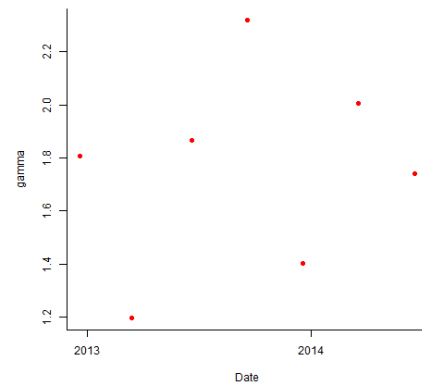
(c) γ_{untrad} , 1-month maturity.



(d) γ_{untrad} , 3-month maturity.



(e) γ_{trad} , 1-month maturity.



(f) γ_{trad} , 3-month maturity.

Figure 7: Relative risk aversion. For the 1 an 3-month maturities we compute $E_t^P(RV^{Var}) = \alpha + \beta \text{Var}_t^{Q_T}(R_{t,T})$ and then γ_{untrad} and γ_{trad} according to Equations (26) and (28).

EVALUATING MODELS JOINTLY WITH ECONOMIC AND STATISTICAL CRITERIA

PAUL SCHNEIDER AND DAVIDE TEDESCHINI

ABSTRACT. We introduce a new criterion for estimation of models used in finance, which explicitly incorporates the models' ability to provide signals for trading strategies. An out-of-sample analysis reveals that an investor using this estimator may enjoy significant excess returns over a competitor who employs purely statistical criteria such as GMM or ML.

1. INTRODUCTION

When designing, selecting, and estimating models for asset allocation, trading signals or risk management a number of simple-to-ask, but hard-to-answer questions arise: Which dynamic trading strategy gives me the highest Sharpe ratio? My model's predictions work well in one trading strategy; will they do well in other strategies? I can predict well according to standard criteria, why am I still losing money when I use my model's signals in a trading strategy? In this paper we show that these questions are related and that they pose an additional layer of complication on top of model selection based on statistical criteria, a hard topic in its own right. In the equity index market, for instance, [Ang and Bekaert \(2007\)](#) and [Welch and Goyal \(2008\)](#) cast doubt on the abilities of elaborate models to beat the historical average in predicting stock returns. [Campbell and Thompson \(2008\)](#), [Cochrane \(2008\)](#), and [Rapach et al. \(2010\)](#) have more encouraging

SWISS FINANCE INSTITUTE, UNIVERSITY OF LUGANO, VIA BUFFI 13, CH-6900 LUGANO.

E-mail address: paul.schneider@usi.ch, davide.tedeschini@usi.ch.

Key words and phrases. Model estimation, pricing kernel, model risk, trading strategy, predictability.

Financial support from SNSF (Projects "Trading Asset Pricing Models" and "Equity risk premia: model evaluation, trading strategies and estimation of implied moments") is gratefully acknowledged.

results, but it is not clear how the predictability, if there was any, would best be translated into a trading strategy and how it would map into gains and losses.

Risk premia can be interpreted as expected profits from trading strategies and there is a growing literature on equity, variance and skew risk premia in the S&P 500 market ([Bakshi et al., 2003](#); [Bondarenko, 2003](#); [Bollerslev et al., 2009](#); [Carr and Wu, 2009](#); [Neuberger, 2012](#)). From economic theory, these risk premia are all functions of the volatility of the pricing kernel and therefore one might ex-ante expect them to co-move. While variance and skew risk premia have shown remarkable similarities in [Kozhan et al. \(2013\)](#), they both seem to be unconnected to the equity premium. This example highlights that assets can be exposed to multiple risk factors and carry multiple risk premia.

To investigate the connection between the different risk compensations we use the likelihood ratio swap, introduced by [Schneider \(2015\)](#). As variance swaps trades implied variance for realized variance, the likelihood ratio swap trades *implied pricing kernel variance* for *realized pricing kernel variance*. An investor can directly trade this instrument. The resulting trading strategy specifies portfolio positions on n-th moment swaps explicitly linked to the model parameters and the option-implied forward-neutral density, thereby relating the predictive density of a model to how much money could be made or lost with it in the market. It thus combines statistical and economic information, addressing the concerns raised in [Leitch and Tanner \(1991\)](#) about the possible divergence of statistical and economic predictability criteria. Moreover, attaining the Hansen Jagannathan bound, the likelihood ratio swap measures exposure to the whole pricing kernel risk.

In the S&P 500 index market we investigate whether a criterion based on optimizing the Sharpe ratio of the likelihood trading strategy improves the predictive power of a financial model. We are particularly interested in the out-of-sample behavior of returns generated by model estimations incorporating economic criteria. To achieve this goal, we test the

out-of-sample performances of the Black-Scholes and of the Levy model based on the homoscedastic Bilateral Gamma distribution from [Küchler and Tappe \(2008\)](#). The out-of-sample measure is used in order to alleviate concerns raised in [Thornton and Valente \(2012\)](#) about the relevance of in-sample economic measures. Then, the profitability of the model is evaluated comparing the likelihood ratio estimation with the Generalized Method of Moments (GMM) of [Hansen and Singleton \(1982\)](#).

The paper is organized as follows. We briefly introduce the likelihood ratio swap in Section 2. Then, Section 3 explains how it can be replicated and traded using a model, how excess returns from this swap are a measure of predictability and concludes with the definition of the “Economic Estimator”. These two sections draw heavily on the results in [Schneider \(2015\)](#). Section 4 is dedicated to the model estimation and Section 5 concludes. Appendix contains proofs for the claims made in the main text.

2. THE LIKELIHOOD RATIO SWAP

In this section we show how the likelihood ratio swap may be used to construct a trading strategy written on a financial model $\mathcal{M}(\theta)$. The simplest trading strategy we can think of consists in entering a forward contract on the asset at time t , and hold it until expiry at time T . Denoting the forward price of an asset contracted at time t with maturity T by $F_{t,T}$ and the forward pricing probability measure by \mathbb{Q}_T , the profit can be written as¹

$$F_{T,T} - \mathbb{E}_t^{\mathbb{Q}_T} [F_{T,T}] = F_{T,T} - F_{t,T}, \quad (1)$$

¹We work with forwards under the T forward measure for ease of exposition, to avoid notational complications through stochastic dividends and interest rates.

and the corresponding (conditional) risk premium is the conditional expectation thereof under the historical, or time-series probability measure \mathbb{P}

$$\mathbb{E}_t^{\mathbb{P}} [F_{T,T}] - \mathbb{E}_t^{\mathbb{Q}^T} [F_{T,T}]. \quad (2)$$

Introducing the likelihood ratio, Radon-Nykodim derivative, or pricing kernel

$$\mathcal{L} := \frac{d\mathbb{Q}_T}{d\mathbb{P}}, \quad (3)$$

the risk premium can be expressed as minus the conditional covariance of the forward price with the pricing kernel: $-Cov_t^{\mathbb{P}}(\mathcal{L}, F_{T,T})$. Another interpretation of the risk premium is as the expected payoff from a swap contract, the notion adopted in this paper.

The main problem associated with definition (2) is that the risk premium depends on the first moment of the forward only. It is therefore necessary to consider other payoff functions. Let

$$R_{t,T} := \log \left(\frac{F_{T,T}}{F_{t,T}} \right),$$

and assume that $\mathcal{L} \in L_{\mathbb{P}}^2$, which guarantees the existence of a finite conditional variance of the pricing kernel.² Then, we define:

$$\mathcal{L}(R_{t,T}) := \mathbb{E}^{\mathbb{P}} [\mathcal{L} \mid R_{t,T}]. \quad (4)$$

This definition creates exposure to the variance of the pricing kernel and projects it onto simple forward returns without assuming that the pricing kernel is spanned by only $R_{t,T}$.

²Define the weighted Hilbert space L_w^2 as the set of (equivalence classes of) measurable functions f on \mathbb{R} with finite L_w^2 -norm defined by

$$\|f\|_{L_w^2}^2 = \int_{\mathbb{R}} |f(\xi)|^2 dw(\xi) < \infty.$$

Accordingly, the scalar product on L_w^2 is denoted by

$$\langle f, h \rangle_{L_w^2} = \int_{\mathbb{R}} f(\xi) h(\xi) dw(\xi).$$

The likelihood ratio swap is simply a swap contract with net payoff $\mathcal{L}(R_{t,T}) - \mathbb{E}_t^{\mathbb{Q}^T} [\mathcal{L}(R_{t,T})]$. Result 2.1 in [Schneider \(2015\)](#) guarantees that its risk premium will be bounded, provided that the tails of the \mathbb{P} distribution are wide enough. Precise technical conditions are discussed in [Filipović et al. \(2013\)](#). In economic terms, these conditions imply that a likelihood ratio contract makes sense only in a world where the investors' fear of severe tail events is occasionally matched or even exceeded by reality, and [Backus et al. \(2011\)](#) find strong evidence for that. At the same time, there must be sufficient heterogeneity in the market, meaning that both insurance sellers and buyers should operate in it. They will take a short and a long position on the likelihood ratio swap respectively.

Unfortunately, the contract introduced above is only a theoretical benchmark claim, since the likelihood ratio is unattainable through a trading strategy. However, the next section shows how it can nevertheless be traded in financial markets and adopted for model estimation purposes.

3. THE ECONOMIC ESTIMATOR

An investor who needs to assess the predictive abilities of a model in economic terms, ideally could write financial contracts on the forecast errors across the entire distribution. In practice, usually, only forecasts of levels are evaluated. A trading strategy based on such a signal would suggest entering into a long forward position on the stock today for an expected profit if the level prediction was higher than the current forward price, and going short otherwise. The net payoff from the corresponding strategy, a first-moment swap, was given in equation [\(1\)](#).

There are many reasons why such an approach can yield surprisingly bad economic results, even with unbiased forecast errors: excess returns could be very small on average, with high variation, as it is the case of the first moment swaps on the S&P 500. Furthermore, neglecting higher-order moments can lead to large unexpected losses. Even

if higher-order moments were taken into account, for instance by simultaneously trading first-moment swaps and variance swaps, it is not clear how one would simultaneously hedge the joint exposure. The reason is that first-moment swaps and second-moment swaps do not move independently, because they are exposed to moments of the same distribution. This can already be seen from the decomposition of the pricing kernel in terms of its cumulants (Backus et al., 2011). To meet these concerns, a trading strategy based on the conditional likelihood ratio in (4) relates the predictive \mathbb{P} density to the market-implied \mathbb{Q}_T density and thereby captures risk premia simultaneously across the entire distribution of returns.

In order to render this strategy operative, let $\mathbb{P}_{\mathcal{M}(\theta)}$, denote the probability measure induced by a model $\mathcal{M}(\theta)$ with parameter θ and introduce³

$$\mathcal{L}_{\mathcal{M}(\theta)} := \frac{d\mathbb{Q}_T}{d\mathbb{P}_{\mathcal{M}(\theta)}}, \quad \text{and} \quad \mathcal{L}_{\mathcal{M}(\theta)}(R_{t,T}) := \mathbb{E}^{\mathbb{P}_{\mathcal{M}(\theta)}} [\mathcal{L}_{\mathcal{M}(\theta)} \mid R_{t,T}]. \quad (5)$$

As long as sufficiently many options are written on $R_{t,T}$, the \mathbb{Q}_T distribution can be computed through the results of Breeden and Litzenberger (1978) and Carr and Madan (2001). Provided that the density function of $\mathbb{P}_{\mathcal{M}(\theta)}$ is known, the likelihood ratio $\mathcal{L}_{\mathcal{M}(\theta)}$ above is fully tractable and, as shown by Schneider (2015), attains the Hansen Jaganatthan bound. Hence, in an ideal arbitrage-free economy where reality is fully captured by the parametric model $\mathcal{M}(\theta)$ ($\mathbb{P}_{\mathcal{M}(\theta)} = \mathbb{P}$), the claim will dominate all the other tradable assets in terms of risk-return ratio. Trading a likelihood ratio swap is therefore equivalent to trade the model $\mathcal{M}(\theta)$. In particular, if $\mathcal{M}(\theta)$ was well-specified, meaning that $\mathbb{P}_{\mathcal{M}(\theta)}$ was close to \mathbb{P} , the Sharpe ratio of $\mathcal{L}_{\mathcal{M}(\theta)}(R_{t,T})$ will be close to the one of the unobservable

³Analogous to the payoff in (4) the list of conditioning arguments could be extended with other asset returns. The payoff above is conditional only on $R_{t,T}$ for ease of exposition and to reflect the availability of option prices. A higher-order expansion in the multivariate case requires options written on joint payoffs. To the best of our knowledge this is only the case in the foreign exchange market, where claims on products of exchange rates can be engineered through the no-arbitrage cross rate condition, and through spread options which are used in the commodity market.

likelihood ratio \mathcal{L} . If not, then the associated trading strategy's Sharpe ratio would be away from the Hansen-Jagannathan bound.

Provided that both $\mathbb{P}_{\mathcal{M}(\theta)}$ and \mathbb{Q}_T have exponential tails and $\mathcal{L}_{\mathcal{M}(\theta)} \in L^2_{\mathbb{P}_{\mathcal{M}(\theta)}}$, Lemmas 3.1 and 3.3 in [Filipović et al. \(2013\)](#) guarantee that the set of polynomials is dense, and that an orthonormal basis of polynomials exists in $L^2_{\mathbb{P}_{\mathcal{M}(\theta)}}$. This admits a polynomial representation

$$\mathcal{L}_{\mathcal{M}(\theta)}(R_{t,T}) = 1 + \sum_{j=1}^{\infty} c_j(\theta, \mathbb{Q}_T) H_j(R_{t,T} \mid \theta), \quad (6)$$

which we truncate for practical purposes

$$\mathcal{L}_{\mathcal{M}(\theta)}^{(J)}(R_{t,T}) := 1 + \sum_{j=1}^J c_j(\theta, \mathbb{Q}_T) H_j(R_{t,T} \mid \theta) \quad (7)$$

$$= \sum_{j=0}^J a_j(\theta, \mathbb{Q}_T) R_{t,T}^j. \quad (8)$$

The orthonormal polynomials $H_j(R_{t,T} \mid \theta)$ are specific to $\mathbb{P}_{\mathcal{M}(\theta)}$ and the algorithm for developing them order by order from the canonical basis can be found in [Appendix A](#). The coefficients $c_j(\theta, \mathbb{Q}_T)$ jointly depend on $\mathbb{P}_{\mathcal{M}(\theta)}$ and \mathbb{Q}_T and they can be obtained in two ways. In the first method \mathbb{Q}_T expectations of $H_j^2(R_{t,T} \mid \theta)$ are computed by applying the [Carr and Madan \(2001\)](#) and [Bakshi and Madan \(2000\)](#) formulas repeatedly. The second method is based on a direct expansion the parametric \mathbb{Q}_T implied by the model. The coefficients $a_j(\theta, \mathbb{Q}_T)$ are defined implicitly by collecting terms of order $R_{t,T}^j$.

Having obtained the truncated representation (8), we can construct a trading strategy for the likelihood ratio swap through the definition of the (n)-th moment swap return over the period (t, T) :

$$Ret_{t,T}^{(n)} := \frac{R_{t,T}^{(n)}}{\mathbb{E}_t^{\mathbb{Q}_T} [R_{t,T}^{(n)}]} - 1, \quad (9)$$

the likelihood ratio excess return is then computed from (8) by trading weighted moment swaps

$$Ret_{t,T}^{\mathcal{L}_{\mathcal{M}(\theta)}(J)}(\theta) := \sum_{n=1}^J a_n(\theta, \mathbb{Q}_T) \cdot Ret_{t,T}^{(n)}, \quad (10)$$

which are a function of the parameters of the model $\mathcal{M}(\theta)$ and of the current option prices. By construction, this portfolio exposes the model to the distribution of forecast errors through the conditional $\mathbb{P}_{\mathcal{M}(\theta)}$ moments and measures this exposure in a model-free way, being $Ret_{t,T}^{(n)}$ model-free.

The characteristics of the described portfolio naturally lead to define a new criterion for model estimation which explicitly measures the predictability of a model in terms of economic value. To this end, consider a generic period of time $[t_0, \dots, tn]$, where we observe the n realizations of the $\mathcal{L}_{\mathcal{M}(\theta)}(R_{t,T})$ net payoffs:

$$Ret_i^{\mathcal{L}_{\mathcal{M}(\theta)}(J)} := \mathcal{L}_{\mathcal{M}(\theta)}(R_{t_i, t_{i+1}}) - \mathbb{E}_{t_i}^{\mathbb{Q}_T} [\mathcal{L}_{\mathcal{M}(\theta)}(R_{t_i, t_{i+1}})], \quad i = 0, 1, \dots, n-1. \quad (11)$$

Then, we define the “Economic Estimator” $\hat{\theta}_{EE}$ as the set of model parameters θ which maximizes the absolute value of the likelihood ratio swap Sharpe ratio, under the constraint of pricing kernel integrating to 1:

$$\begin{aligned} \hat{\theta}_{EE} := \arg\max_{\theta} & \quad \frac{\left| \frac{1}{n} \sum_{i=1}^n Ret_i^{\mathcal{L}_{\mathcal{M}(\theta)}(J)} \right|}{\text{std} \left(Ret_i^{\mathcal{L}_{\mathcal{M}(\theta)}(J)} \right)} \\ \text{s.t.} & \quad \frac{1}{n} \sum_{i=1}^n \mathcal{L}_{\mathcal{M}(\theta)}^{(J)}(R_{t_i, t_{i+1}}) = 1. \end{aligned} \quad (12)$$

Noticeably, the estimator is constructed in order to have a direct link between the model parameters, which appear in the weights $a_n(\theta, \mathbb{Q}_T)$ and the criterion to optimize. This direct relation is obtained in the same spirit of [Brandt \(1999\)](#) and [Aït-Sahalia and Brandt \(2001\)](#), but outside of the GMM and Conditional GMM framework.

The Economic Estimator definition is justified by the following result:

Result 3.1. *In a complete arbitrage-free market, where an asset whose dynamics is driven by the set of parameters $\theta_{\mathbb{P}_{\mathcal{M}(\theta)}}$ is traded, the optimal point is unique and is given by $\hat{\theta}_{EE} = \theta_{\mathbb{P}_{\mathcal{M}(\theta)}}$*

Proof. In complete markets there exists only one $\mathcal{L}_\theta(R_{t,T})$ satisfying the constraint in (12) because the stochastic discount factor is unique. Since $\mathcal{L}_\theta(R_{t,T})$ also achieves the highest Sharpe ratio in the market, the only point such that (12) is maximized should be $\hat{\theta}_{EE} = \theta_{\mathbb{P}_{\mathcal{M}(\theta)}}$. \square

4. EMPIRICS

The ability of a financial model to generate profitable trading strategies when the parameters are obtained through the Economic Estimator is tested in the S&P 500 market. Black-Scholes (BS) and Bilateral Gamma (BL) model from Küchler and Tappe (2008) are examined.

4.1. Data and statistical properties. The S&P 500 and options data are taken from OptionMetrics and include closing bid and ask quotes for each option contract along with the corresponding strike price, Black-Scholes implied volatility, the zero-yield curve, and dividend yield. From the data we filter out all entries with non-standard settlements and with implied volatility less than 0.001 or higher than 9. The options mature every third Friday each month, and we use this maturity in a monthly time grid. Joint option and S&P 500 data is available from March 1996 up to January 2015. We use the sample period from March 1996 to December 2012 as a burn-in period and the remainder for the out-of-sample study.

With the yield curve and dividend yield information from OptionMetrics we construct a time series of forwards on the S&P 500 spot index. To estimate the conditional \mathbb{Q}_T

moments for the trading strategy from Section 3, we apply the formula from Carr and Madan (2001) to log-forward returns.

4.2. The Black Scholes model as robustness check. In order to assess the robustness of our methodology, we test the optimization procedure illustrated in the previous section within the Black-Scholes framework. Indeed, in this particular case, the shape of the pricing kernel is known and there is no need to implement the polynomial expansion for approximating its value. Hence, this model permits to compare the parameters resulting from optimizing the polynomial expansion with the ones obtained by considering directly the pricing kernel, which will represent our benchmark case. More specifically, we assume the \mathbb{P} distribution of log returns to be $N(\mu - \frac{1}{2}\sigma^2 T, \sigma^2 T)$ and the \mathbb{Q}_T distribution to be $N(-\frac{1}{2}\sigma^2 T, \sigma^2 T)$, since the forward price is a \mathbb{Q}_T -martingale. As shown in Schneider (2015), the likelihood ratio projection has a known form:

$$\mathcal{L}_{BS}(R_{t,T}) = \exp(\alpha + \beta R_{t,T}), \quad (13)$$

where $\alpha = \frac{\mu T(\mu - \sigma^2)}{2\sigma^2}$ and $\beta = -\mu/\sigma^2$. The explicit expression for the pricing kernel allows to compare the results obtained by adopting the polynomial expansion (denoted as $\{\hat{\mu}^J, \hat{\sigma}^J\}$ hereafter), with the ones retrieved by optimizing directly the expression in (13):

$$\begin{aligned} \{\hat{\mu}, \hat{\sigma}\}_{EE} := & \underset{\theta}{\operatorname{argmax}} \left(\frac{|\frac{1}{n} \sum_{i=1}^n Ret_i^{\mathcal{L}_{BS}}|}{\operatorname{std}(Ret_i^{\mathcal{L}_{BS}})} \right) \\ \text{s.t. } & \frac{1}{n} \sum_{i=1}^n \mathcal{L}_{BS}(R_{t_i, t_{i+1}}) = 1, \end{aligned} \quad (14)$$

where

$$Ret_i^{\mathcal{L}_{BS}} := \mathcal{L}_{BS}(R_{t_i, t_{i+1}}) - \mathbb{E}_t^{\mathbb{Q}_T} [\mathcal{L}_{BS}(R_{t_i, t_{i+1}})], \quad i = 0, 1, \dots, n-1. \quad (15)$$

Table 1 displays the optimized parameters. Being the two sets of estimates very close, we can deduce that the intrinsic error of the approximation does not substantially affect the maximization procedure.

	$\hat{\mu}$	$\hat{\mu}^J$	$\hat{\sigma}$	$\hat{\sigma}^J$
Mean	0.110	0.119	0.194	0.206
Standard deviation	0.123	0.133	0.177	0.137

TABLE 1. **Black Scholes parameters summary statistics:** the table reports mean and standard deviation of μ and σ computed through the Economic Estimator. For $\hat{\mu}$ and $\hat{\sigma}$ the explicit expression of the pricing kernel is involved in the optimization, whereas to estimate $\hat{\mu}^J$ and $\hat{\sigma}^J$ the polynomial expansion is used.

4.3. **Out of sample.** In this section we compare the out-of-sample performances of the Black-Scholes and of the homoscedastic Bilateral Gamma model for log S&P 500 forward returns, when two competing estimation criteria are used: the Economic Estimator defined in (3) and the Generalized Method of Moments (GMM). The Bilateral Gamma model does not accommodate stochastic volatility, but the distribution of simple index returns can be parameterized to exhibit flexible skewness and sizable excess kurtosis. Moreover, the tails of the Bilateral Gamma distributions are heavy enough to render expansion (6) meaningful.

Starting from the initial burn-in sample we estimate the Bilateral Gamma model with the two methodologies. For the Economic Estimator, we use a fourth-order likelihood expansion from Filipović et al. (2013). In the Black-Scholes case μ and σ are computed through sample moments, while the Bilateral Gamma parameters can be retrieved by solving the system of equations given in Küchler and Tappe (2008) and then calibrating the model with call options.

The first estimation period is $[t_0 = 03/14/1996, t_n = 12/19/2002]$ produces the first two sets of parameters. Collecting terms and using the conditional \mathbb{Q}_T moments estimated from option prices again, we then compute the fixed leg $\mathbb{E}_{t_0}^{\mathbb{Q}_T} \left[\mathcal{L}_{\mathcal{M}(\theta)}^{(4)}(R_{t_0, t_1}) \right]$ of the likelihood ratio swap written on model $\mathcal{M}(\theta)$. The differences between the realized counterparts developed in Section 3 and the fixed legs give the excess returns on the likelihood ratio swap approximation, where the weights are calculated with the parameters estimated in t_0 . We then move on to estimate the models again including the data point at time $t_1 = 04/16/2009$, compute the returns at time $t_2 = 05/14/2009$ and so forth. The payoffs used in this exercise are therefore entirely out-of-sample. Summary statistics on the Bilateral Gamma parameters resulting from the estimation are contained in Table 2.

	Mean(EE)	Standard deviation(EE)	Mean(GMM)	Standard deviation(GMM)
$\hat{\alpha}^+$	3.538	3.462	5.180	2.865
$\hat{\alpha}^-$	5.002	3.235	1.240	0.518
$\hat{\lambda}_{\mathbb{P}}^+$	62.950	72.430	94.205	31.230
$\hat{\lambda}_{\mathbb{P}}^-$	97.152	77.712	25.292	6.182
$\hat{\lambda}_{\mathbb{Q}_T}^+$	121.7276	99.852	166.472	87.026
$\hat{\lambda}_{\mathbb{Q}_T}^-$	180.260	96.592	40.166	20.502

TABLE 2. **Bilateral Gamma parameters summary statistics:** mean and standard deviation of the six Bilateral Gamma parameters estimated with the Economic Estimator and with GMM. More specifically, $\hat{\alpha}^+$, $\hat{\alpha}^-$, $\hat{\lambda}_{\mathbb{P}}^+$ and $\hat{\lambda}_{\mathbb{P}}^-$ are obtained by matching the first four moments of the S&P500 return distribution. Given the values of these parameters, $\hat{\lambda}_{\mathbb{Q}_T}^+$ and $\hat{\lambda}_{\mathbb{Q}_T}^-$ are estimated with a standard option calibration. A detailed explanation of the methodology can be found in Küchler and Tappe (2008).

Figure 1 shows the returns from the Black-Scholes and Bilateral Gamma likelihood ratio swaps, when the parameters are estimated. Concerning the first model, the time series co-move almost one-to-one, whereas the Bilateral Gamma returns differ dramatically, in particular during crises. However, in both cases the investor adopting the Economic Estimator can generate large negative returns, indicating a significant advantage over the GMM.

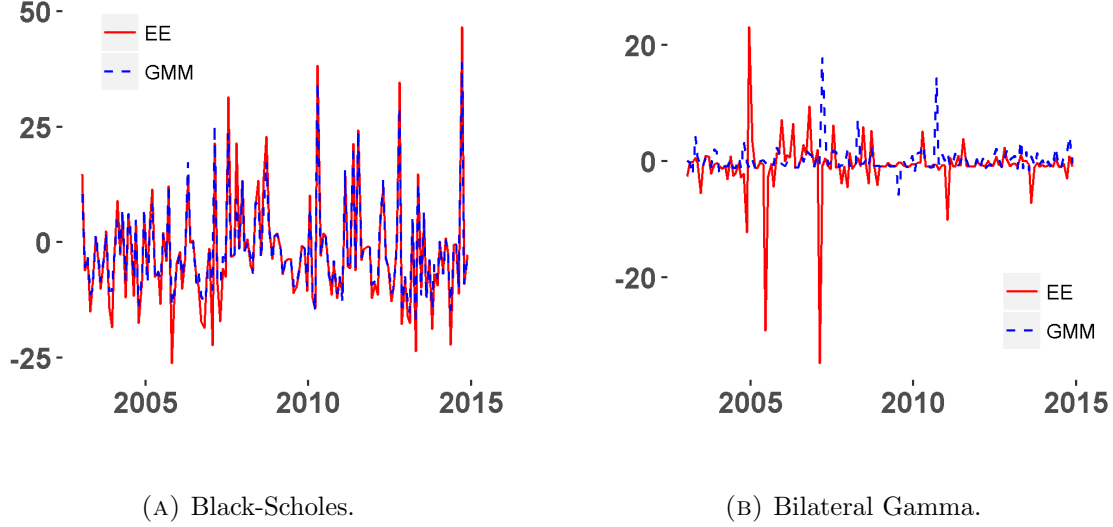


FIGURE 1. **Out of sample returns:** The figure shows monthly out-of sample returns on the likelihood ratio swap trading strategy from 10, where the weights $a_n(\theta, \mathbb{Q}_T)$ are based on two different sets of parameters (Economic Estimator and GMM) and on two different financial models (Black-Scholes and Bilateral Gamma model). The time series are computed from January 2003 until January 2015 after a burn-in sample going from March 1996 to December 2002.

The return time series are serially uncorrelated, allowing an independence bootstrap. Since the likelihood ratio swap can be interpreted as a hedging asset, the highest possible Sharpe Ratio is achieved through a short position. Section 4.3 shows the sampling distribution of the model-implied Sharpe ratios of the likelihood-ratio strategy. The result suggests that, according to the Sharpe ratio criterion, the Economic Estimator models outperforms GMM in most states of the world. The point estimates of the GMM and Economic Estimator strategies are monthly Sharpe ratios of -0.18 and -0.20 for the Black-Scholes model, and 0.09 and -0.15 for the Bilateral Gamma model. This finding does not question the importance of statistical methods for financial model estimation, but the inferior return-risk ratio when implemented in a trading strategy. Given the confirmation of the Leitch and Tanner (1991) result of the possible divergence between statistical and economic criteria also for the S&P 500, the question remains how to weight

the conflicting statistical and economic evidence collected so far. How severe is a failure in predicting skewness compared to a failure to predict variance and how does this translate into excess returns from trading strategies? The Economic Estimator helps addressing these questions along two dimensions. First, the likelihood ratio trading rule is based on both the \mathbb{P} density of the model as well as the \mathbb{Q}_T density. The portfolio decision imposed by the likelihood ratio swap thereby takes into account both predictability as well as profitability. Second, the likelihood is benchmarked against the Hansen-Jagannathan bound through the Sharpe ratio, and thus not require the specification of a separate loss function.

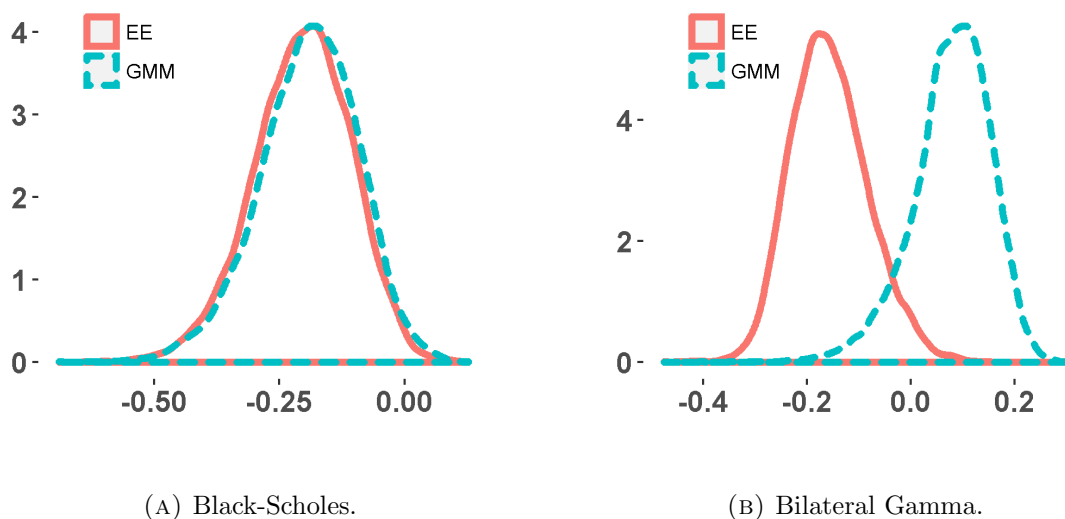


FIGURE 2. **Bootstrap distribution of monthly Sharpe ratios.** The graph shows the bootstrap distribution of monthly out-of-sample Sharpe ratios estimated from the likelihood ratio swap trading strategy on the Black-Scholes model, as well as the Bilateral Gamma model from Küchler and Tappe (2008), with parameters estimated both with the Economic Estimator and the GMM. The time series are computed from January 2003 until January 2015 after a burn-in sample going from March 1996 to December 2002.

5. CONCLUSIONS

In this paper we introduce a framework where financial model estimates are based on the performance of a trading strategy. In particular, the features of the likelihood ratio swap of [Schneider \(2015\)](#) are exploited to construct the Economic Estimator. The likelihood ratio swap attains the Hansen-Jagannathan bound, and can therefore be used as a benchmark instrument. In addition, it is fully tradable using forward and options and can be expressed as a weighted sum of moment swap returns, where the weights depend directly on model parameters.

The Economic Estimator is presented in order to overcome the flaws of statistical criteria shown in [Leitch and Tanner \(1991\)](#). The parameter estimates are retrieved without maximizing a likelihood or imposing moment conditions. Conversely, they are obtained by focusing on the economic value of predictability through the optimization of a trading strategy written on the likelihood ratio swap. The comparison between the Economic Estimator and the GMM estimation, implemented for the Black-Scholes model and for a Levy process based on the Bilateral Gamma distribution, supports the idea to take profitability into account for estimation purposes.

In summary, it is not clear why an agent should rely on standard inference methods to estimate models for investment purposes. Future research will focus on finding further evidence of the better performance of the Economic Estimator, by including in the comparison other financial models, such as stochastic volatility ones.

REFERENCES

Aït-Sahalia, Y. and M. W. Brandt (2001). Variable selection for portfolio choice. *The Journal of Finance* 56(4), 1297–1351.

- Ang, A. and G. Bekaert (2007). Stock return predictability: Is it there? *Review of Financial Studies* 20(3), 651–707.
- Backus, D., M. Chernov, and I. Martin (2011). Disasters implied by equity index options. *The journal of finance* 66(6), 1969–2012.
- Bakshi, G., N. Kapadia, and D. Madan (2003). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies* 16(1), 101–143.
- Bakshi, G. and D. Madan (2000). Spanning and derivative-security valuation. *Journal of Financial Economics* 55(2), 205–238.
- Bollerslev, T., G. Tauchen, and H. Zhou (2009). Expected stock returns and variance risk premia. *Review of Financial Studies* 22(11), 4463–4492.
- Bondarenko, O. (2003). Why are put options so expensive. *University of Illinois at*.
- Brandt, M. W. (1999). Estimating portfolio and consumption choice: A conditional euler equations approach. *The Journal of Finance* 54(5), 1609–1645.
- Breeden, D. T. and R. H. Litzenberger (1978). Prices of state-contingent claims implicit in option prices. *Journal of business*, 621–651.
- Campbell, J. Y. and S. B. Thompson (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies* 21(4), 1509–1531.
- Carr, P. and D. Madan (2001). Towards a theory of volatility trading. *Option Pricing, Interest Rates and Risk Management, Handbooks in Mathematical Finance*, 458–476.
- Carr, P. and L. Wu (2009). Variance risk premiums. *Review of Financial Studies* 22(3), 1311–1341.
- Cochrane, J. H. (2008). The dog that did not bark: A defense of return predictability. *Review of Financial Studies* 21(4), 1533–1575.

- Filipović, D., E. Mayerhofer, and P. Schneider (2013). Density approximations for multivariate affine jump-diffusion processes. *Journal of Econometrics* 176(2), 93–111.
- Hansen, L. P. and K. J. Singleton (1982). Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica: Journal of the Econometric Society*, 1269–1286.
- Kozhan, R., A. Neuberger, and P. Schneider (2013). The skew risk premium in the equity index market. *Review of Financial Studies* 26(9), 2174–2203.
- Küchler, U. and S. Tappe (2008). Bilateral gamma distributions and processes in financial mathematics. *Stochastic Processes and their Applications* 118(2), 261–283.
- Leitch, G. and J. E. Tanner (1991). Economic forecast evaluation: profits versus the conventional error measures. *The American Economic Review*, 580–590.
- Neuberger, A. (2012). Realized skewness. *Review of Financial Studies* 25(11), 3423–3455.
- Rapach, D. E., J. K. Strauss, and G. Zhou (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *Review of Financial Studies* 23(2), 821–862.
- Schneider, P. (2015). Generalized risk premia. *Journal of Financial Economics* 116(3), 487–504.
- Thornton, D. L. and G. Valente (2012). Out-of-sample predictions of bond excess returns and forward rates: An asset allocation perspective. *Review of Financial Studies* 25(10), 3141–3168.
- Welch, I. and A. Goyal (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies* 21(4), 1455–1508.

APPENDIX A. EXPANSION OF LIKELIHOOD RATIO

By existence of exponential moments of the \mathbb{Q}_T and the $\mathbb{P}_{\mathcal{M}(\theta)}$ and technical conditions on the tails of $\mathbb{P}_{\mathcal{M}(\theta)}$ there exists an orthonormal basis of $L^2_{\mathbb{P}_{\mathcal{M}(\theta)}}$. To compute the basis we can employ the Gram-Schmidt process reviewed below.

A.1. $\mathbb{P}_{\mathcal{M}(\theta)}$ Orthonormal Polynomials. The orthonormal polynomials H from 7 can be computed with

Algorithm A.1 (Gram-Schmidt Process).

$$\begin{aligned} H_0(x \mid \theta) &= 1, \\ \tilde{H}_i(x \mid \theta) &= x^i - \sum_{j=0}^{i-1} \int_{\mathbb{R}} \xi^i H_j(\xi \mid \theta) d\mathbb{P}_{\mathcal{M}(\theta)}(\xi \mid \theta) \cdot H_j(x \mid \theta), \\ H_i(x \mid \theta) &= \frac{\tilde{H}_i(x \mid \theta)}{\sqrt{\int_{\mathbb{R}} \tilde{H}_i^2(\xi \mid \theta) d\mathbb{P}_{\mathcal{M}(\theta)}(\xi \mid \theta)}} \end{aligned}$$

Denoting the moments of $\mathbb{P}_{\mathcal{M}(\theta)}$ generically by μ

A.2. Expansion Coefficients. The coefficients in the expansion 7 are obtained through the formula

$$c_i(\theta, \mathbb{Q}_T) = \frac{\mathbb{E}_t^{\mathbb{Q}_T} [\tilde{H}_i(R_{t,T} \mid \theta)]}{\sqrt{\int_{\mathbb{R}} \tilde{H}_i^2(\xi \mid \theta) d\mathbb{P}_{\mathcal{M}(\theta)}(\xi \mid \theta)}}, \quad (16)$$

where for twice-differentiable f we have from Carr and Madan (2001)

$$\mathbb{E}_t^{\mathbb{Q}_T} [f(F_{T,T})] = f(F_{t,T}) + \frac{1}{p_{t,T}} \left(\int_0^{F_{t,T}} f''(K) P_{t,T}(K) dK + \int_{F_{t,T}}^{\infty} f''(K) C_{t,T}(K) dK \right). \quad (17)$$

where $C_{t,T}(K)$ and $P_{t,T}(K)$ denote European Calls and Puts written on the spot underlying at time t with maturity T and strike price K .

APPENDIX B. FIGURES

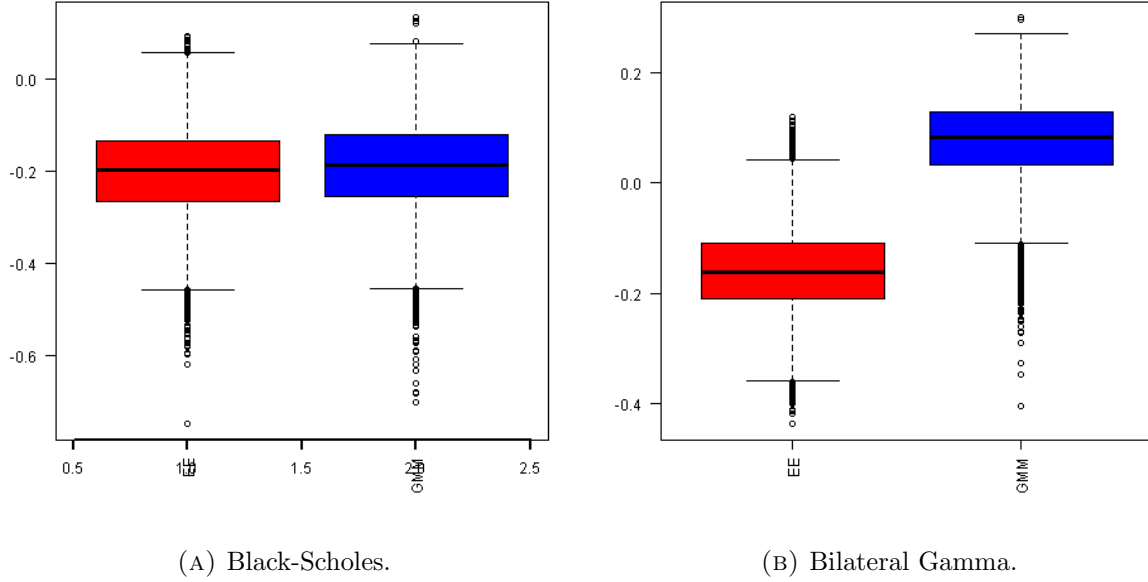


FIGURE 3. **Bootstrap distribution of monthly Sharpe ratios (boxplot).** The boxplot shows the bootstrap distribution of monthly out-of-sample Sharpe ratios estimated from the likelihood ratio swap trading strategy on the Black-Scholes model, as well as the Bilateral Gamma model from [Kühler and Tappe \(2008\)](#), with parameters estimated both with the Economic Estimator and the GMM. The time series are computed from January 2003 until January 2015 after a burn-in sample going from March 1996 to December 2002.