

Recovering the Market Risk Premium from Stock and Option Prices*

George Chalamandaris[†] and Leonidas S. Rompolis[‡]

November 11, 2015

Abstract

We find that the estimation of the market expected return benefits significantly from enriching the information set we use, with increasingly finer detail regarding the shape of the physical and risk-neutral distributions and the shape discrepancy between the two. Assuming the existence of a monotonic projected pricing kernel we extend Duan and Zhang's (2014) theoretical model to a general system of equations that relates physical cumulants of any order to risk-neutral ones through the projected relative risk aversion coefficient. Using stock and option data from the S&P 500 index we employ our general specification to estimate the ex-ante market risk premium for the period 2001-2010 in a monthly frequency. The empirical results strongly support our hypothesis both in a statistical sense and in the context of the present value identity that associates dividend-price ratio to expected returns and dividend growth rates.

Keywords: Ex-ante market risk premium, risk aversion coefficient, physical cumulants, risk-neutral cumulants.

JEL: G12, G17, C51, C53.

*The authors would like to thank George Skiadas for excellent research assistance. We also thank Elias Tzavalis, Alex Taylor, Ser-Huang Poon, seminar participants at Manchester Business School and participants of the 2013 National Conference of FEBS for useful comments and suggestions. Financial support from Research Center of Athens University of Economics and Business (EP-1689-01, EP-1787-01 and EP-2237-01) is gratefully acknowledged.

[†]Department of Accounting and Finance, Athens University of Economics and Business. 76 Patission street, 10434 Athens, Greece. Tel: +30-2108203392. Fax: +30-2108228816. Email: gchalamandar@aueb.gr.

[‡]Corresponding author: Department of Accounting and Finance, Athens University of Economics and Business. 76 Patission street, 10434 Athens, Greece. Tel: +30-2108203465. Fax: +30-2108228816. Email: rompolis@aueb.gr.

1 Introduction

The expected return of an asset (denoted as ER henceforth) is one of the most important concepts in the financial literature, as it constitutes a central input in asset pricing and in risk-return decision making. By its very definition, it is forward-looking in nature, it depends on the systematic risk of the asset and it is practically unobservable.

The most common method for estimating the ER and its associated risk premium amounts to calculating the average value of historical realized returns using a long estimation period. However, this approach is problematic for at least three reasons: First, the average realized return is an unconditional estimate, that becomes equal to its conditional counterpart only under the assumption of i.i.d. innovations. Given that a number of studies document ERs being time-varying and persistent (see Cochrane (2011), *inter alia*), it follows that the conditional ER is not very likely to coincide with the ex-post unconditional one. Second, this ex-post long-run estimate does not take into account short-term changes in market conditions (see Merton (1980)). Third, it is highly probable that the risk premium, being an ex-ante quantity, captures a level of risk which is related to the occurrence of some bad states of nature that will not be realized in the sample after all (the well-known peso problem) (see Brown, Goetzmann and Ross (1995), *inter alia*).

To overcome these deficiencies, three alternative routes were proposed in the literature. The first uses survey on academics, investors or business managers to get their view on the ER (see Welch (2000) and Graham and Harvey (2007)). The second employs the present value identity to estimate the implied cost of equity capital from current stock prices and analysts forecasts for future cash flows (see Pastor, Sinha and Swaminathan (2008)). Both these methodologies are subject to a number of limitations that are qualitatively similar. On one hand, surveys are expressions of personal opinion and as such, they suffer from sample selection bias that is difficult to quantify. On the other hand, analyst forecasts can also be prone to biases, while the implied cost of capital does not generally coincide with the one-period ER, even though it conveys information comparable to it (see Chen, Da and Zhao (2013)).

This leaves a third route for estimating ERs that combines information from the stock and options markets, along with either a parametric option pricing model (see Santa-Clara and Yan (2010)) or a semi-parametric procedure (see Duan and Zhang (2014)). In the context of this third approach, this paper provides new evidence suggesting that the estimation procedure of the ER benefits significantly - in many ways - once we enrich the information set we use with increasingly finer detail regarding the shape of the physical and risk-neutral distributions and the shape discrepancy between the two. Our setup leads us to codify the shape and the discrepancy of the two distributions by means of their respective cumulants.

We build our theoretical framework assuming the existence of a monotonic projected pricing kernel motivated by the well-known power/logarithmic utility function. As our starting point, we extend the theoretical model of Duan and Zhang (2014) to a general system of equations that relates

cumulants of the physical distribution of the market portfolio log-returns (designated henceforth as physical cumulants) of any order to those of the risk-neutral distribution (denoted as risk-neutral cumulants) through the projected relative risk aversion coefficient (denoted as PRRAC henceforth).

The implications of this relation are twofold. First, the ER of the market portfolio log-return distribution, and its associated market risk premium (denoted as MRP henceforth), is connected to higher-order physical cumulants and the PRRAC. That is, investors require a higher compensation to hold the market portfolio, the more their aversion towards risk increase and/or the more their expectations about future returns are characterized by higher level of volatility, more negative skewness and higher excess kurtosis. Second, and most important, it restricts the shape discrepancy between the physical and risk-neutral distributions by means of the PRRAC. This last model implication is very intuitive: If we want to infer the PRRAC from the physical and risk-neutral distributions, it is expedient to use as much information as we can about their shape discrepancy.

Our empirical findings are strongly supportive of our theoretical framework. Qualitatively speaking, the estimated MRP seems sensible, being always positive, time-varying, persistent and counter-cyclical with its annualized values ranging between less than 1% and around 100%. As expected it increases during the two financial turmoil periods (the burst of the Dot-Com bubble in 2002 and the sub-prime mortgage crisis from 2007 to 2009), reaching its highest peak at the aftermath of Lehman's debacle. Moreover, we produce conditional ERs of which the average value is significantly higher (about 17.5%) than the unconditional ex-post one (around 8%). This difference is in line with the peso problem explanation, stating that the average ex-ante MRP should exceed the average ex-post one, simply because the former prices the fear of possible catastrophic states of the economy that cannot be traced in the latter.

Statistically, our analysis provides strong evidence that when we allow our econometric specification to contain higher-order cumulants (thus providing a finer description of the shape of the two distributions) and/or when we impose a larger number of restrictions about the shape discrepancy between the physical and the risk-neutral distributions, the GMM estimator of the PRRAC becomes more efficient.

Moreover, this paper suggests a new approach to assess the plausibility of competing ER estimates in the context of the present value identity. In that exercise, we find that only the ERs produced from the augmented information sets, are consistent with the predictions of the theory associating dividend-price ratio to ERs and dividend growth rates. This step is very important because it confirms our results, based on a framework that is independent from the pricing kernel assumption we have made, while it also employs a set of data that has not been used in the estimation of ERs. Among our other contributions, there is distinct novelty in our testing methodology for this particular step, since it constitutes the first time (to the best of our knowledge) that the relationship between observed dividend-price ratio and ERs is directly examined. In the literature, similar tests rely mostly on the traditional predictive regression approach, which is known to suffer from various shortcomings (see Kojien and Van Nieuwerburgh (2011) for a review).

Our theoretical approach is closely related to previous studies that also used stock and option data to estimate the ex-ante MRP. In contrast to Santa-Clara and Yan (2010) our methodology does not rely on any parametric option pricing model. Compared to Duan and Zhang’s (2014) paper, while our system specification is based on the same theoretical premises (i.e., the same pricing kernel), it does extend their setup to a larger number of equations and cumulant restrictions. This enables us to produce more accurate estimates of the PRRAC which, in turn, generate ER estimates that satisfy almost perfectly the present value identity.

The paper is organized as follows. Section 2 derives the theoretical model. Section 3 presents the econometric formulation for estimating market ERs and the associated ex-ante MRP. Section 4 describes the data and discuss the estimates of the MRP. Section 5 compares the ER time series estimates on the basis of the present value identity. Section 6 concludes the paper. All of the derivations are given in a technical appendix.

2 Theoretical model

Consider an arbitrage-free economy with a pricing kernel $M_t(r_{t,T})$ projected onto the space of the market portfolio log-return $r_{t,T} \equiv \ln(S_T/S_t)$, where S_T (or S_t) denotes the market portfolio price at the future time T (or at the current time t).¹ Then, the conditional on the current market information set \mathcal{I}_t distribution of $r_{t,T}$ under the risk-neutral measure Q , denoted as $f_t^Q(r_{t,T})$, is given as:

$$f_t^Q(r_{t,T}) = e^{r_f \tau} M_t(r_{t,T}) f_t^P(r_{t,T}). \quad (1)$$

In formula (1), $f_t^P(r_{t,T})$ is the conditional distribution of $r_{t,T}$ under the physical (objective) measure P , r_f is the annualized continuously compounded risk-free rate and $\tau = T - t$. If we further assume that $M_t(r_{t,T})$ is implied by a power/logarithmic utility function, i.e., $M_t(r_{t,T}) = \beta e^{-\gamma r_{t,T}}$, where β is a scaling factor and γ is the PRRAC,² it follows that we can express formula (1) in terms of moment-generating functions as:

$$m_t^Q(u) = \ln \beta + r_f \tau + m_t^P(u - \gamma). \quad (2)$$

In formula (2), $m_t^P(u) = \ln E_t^P [e^{ur_{t,T}}]$ and $m_t^Q(u) = \ln E_t^Q [e^{ur_{t,T}}]$ stand for the logarithms of the conditional moment-generating function of $r_{t,T}$ under measures P and Q , respectively. Based on the previous relation, we can deduce a system of equations that relates the two measures in terms

¹This pricing kernel is known as the projected pricing kernel (see Rosenberg and Engle (2002)). Under the assumptions that the asset level is equal to the aggregate wealth and that investors have a finite horizon, the projected pricing kernel is equal to the original pricing kernel (see, e.g. Ait-Sahalia and Lo (2000)).

²The projected pricing kernel implied by a power/logarithmic utility has been employed in many recent studies of the literature to estimate the risk aversion coefficient and the log-return distribution under measure P , without specifying any state variable (see Rosenberg and Engle (2002), Bakshi, Kapadia and Madan (2003) and Bliss and Panigirtzoglou (2004), *inter alia*).

of their cumulants. This is given in the following Proposition.

Proposition 1 *Let $k_{t,n}^Q$ and $k_{t,n}^P$ be the n^{th} -order cumulants of the τ -period log-return $r_{t,T}$ distribution conditional on the current market information set \mathcal{I}_t under the physical P and risk-neutral Q measure, respectively. Then, the following relation holds:*

$$k_{t,n}^Q = \sum_{m=0}^{\infty} k_{t,n+m}^P \frac{(-\gamma)^m}{m!}, \quad \forall n \in \mathbb{N}. \quad (3)$$

Individual approximate equations of this system has been proved in the literature during the previous years. In particular, Bakshi, Kapadia and Madan (2003), Bakshi and Madan (2006) and Duan and Zhang (2014) give approximate counterparts of (3) based on variance, skewness and kurtosis coefficients for $n = 3$, $n = 2$ and $n = 1$, respectively. Equation (3) is an exact formula that can be employed for all n . In that context, Proposition 1 can be seen as refining and generalizing existing results in the literature.

Note also here that there exist an equivalent, to (3), formula that relates physical cumulants to risk-neutral ones through the PRRAC (see Rompolis and Tzavalis (2010)). This is given as follows:

$$k_{t,n}^P = \sum_{m=0}^{\infty} k_{t,n+m}^Q \frac{\gamma^m}{m!}, \quad \forall n \in \mathbb{N}. \quad (4)$$

Formulas (3) and (4) provides a unified framework that restricts the risk-neutral (or physical) distribution by taking its physical (or risk-neutral) counterpart and γ as given. Most importantly, it associates the premium that investors are requiring so as to take on risk with higher-order physical (or risk-neutral) cumulants and the PRRAC. If the latter is unknown and we know instead the cumulants of both distributions, then we can use a set of equations implied by formula (3) or (4) first to estimate γ and then to retrieve the MRP.

For $n = 1$ formula (3) implies that:

$$k_{t,1}^P - k_{t,1}^Q = \gamma k_{t,2}^P - \frac{\gamma^2}{2!} k_{t,3}^P + \frac{\gamma^3}{3!} k_{t,4}^P + \dots \quad (5)$$

This clearly indicates that the MRP, defined as $k_{t,1}^P - k_{t,1}^Q$, is positively related to the physical conditional variance, $k_{t,2}^P$, and excess kurtosis, $k_{t,4}^P$ and negatively related to the skewness of the distribution, measured by $k_{t,3}^P$.³ The result is intuitive, since it stipulates that the MRP increases when the likely losses become increasingly more severe. From the viewpoint of portfolio management, this direct association of ER $k_{t,1}^P$ to higher-order physical cumulants is supportive of recent advances in portfolio optimization that detract from the Markowitz paradigm (see Goh, Lim, Sim

³More precisely we can write the skewness and kurtosis coefficients as $Sk_t^P = \frac{k_{t,3}^P}{(k_{t,2}^P)^{3/2}}$ and $Ku_t^P = 3 + \frac{k_{t,4}^P}{(k_{t,2}^P)^2}$, respectively.

and Zhang (2012), *inter alia*).⁴ Formula (5) also indicates that if the log-return $r_{t,T}$ follows the normal distribution, which implies that $k_{t,n}^P = 0$ for $n > 2$, then the MRP can be written as:

$$k_{t,1}^P - k_{t,1}^Q = \gamma k_{t,2}^P, \quad (6)$$

which is a well-known result implied by the CAPM and the Black-Scholes model assumptions.

Formula (5) indicates that knowledge of γ and of higher-order physical cumulants provides an estimate of the ex-ante MRP. Note here that evaluating formula (4) for $n = 1$ gives an equivalent way to estimate the same quantity, if we take as given the PRRAC and risk-neutral cumulants:

$$k_{t,1}^P - k_{t,1}^Q = \gamma k_{t,2}^Q + \frac{\gamma^2}{2!} k_{t,3}^Q + \frac{\gamma^3}{3!} k_{t,4}^Q + \dots \quad (7)$$

In the empirical part of the paper, we will apply alternatively formulas (5) and (7) so as to calculate the ex-ante MRP.

To estimate the PRRAC γ , we employ the information contained in the rest of the equations in (3) or (4), which correspond to higher-order cumulants. The intuition behind this step is very interesting. By rewriting each equation in (3) or (4) for $n \geq 2$, in terms of difference $k_{t,n}^P - k_{t,n}^Q$, we retrieve a system of equations restricting the premia between the cumulants of the same order. Specifically, for $n = 2$ formulas (3) and (4) lead to two equivalent expressions for the variance risk premium,

$$k_{t,2}^P - k_{t,2}^Q = \gamma k_{t,3}^P - \frac{\gamma^2}{2!} k_{t,4}^P + \dots \quad (8)$$

and

$$k_{t,2}^P - k_{t,2}^Q = \gamma k_{t,3}^Q + \frac{\gamma^2}{2!} k_{t,4}^Q + \dots, \quad (9)$$

⁴Note here that Duan and Zhang (2014) define the MRP as $k_{t,1}^P - (r_f - \delta)\tau$, where δ denotes the annualized continuously compounded dividend yield. However, since $k_{t,1}^P$ is the mean of the log-return distribution we argue that the corresponding moment under risk-neutral measure Q , $k_{t,1}^Q$, should be subtracted in order to appropriately define the MRP. The latter is the expected continuously compounded return under Q . In contrast, $r_f - \delta$ is the annualized expected percentage price change which occurs in a very short period of time.

Given that the risk-neutral mean can be written as:

$$k_{t,1}^Q = (r_f - \delta)\tau - \sum_{m=2}^{\infty} \frac{k_{t,m}^Q}{m!}$$

(see Chalamandaris and Rompolis (2012) for the proof), the ex-ante MRP is equal to:

$$k_{t,1}^P - k_{t,1}^Q = k_{t,1}^P - (r_f - \delta)\tau + \sum_{m=2}^{\infty} \frac{k_{t,m}^Q}{m!}.$$

Due to Jensen's inequality, it can be easily shown that $k_{t,1}^Q < (r_f - \delta)\tau$, implying that Duan and Zhang's (2014) estimates of the MRP will be consistently lower than ours.

Moreover, the definition of Duan and Zhang (2014) for the MRP implies the rather uncomfortable result that this is different from zero when investors are risk-neutral, i.e., $\gamma = 0$.

respectively. Equation (8) indicates that the empirically documented difference between physical and risk-neutral variance (see Christensen and Prabhala (1998), *inter alia*) can be attributed to higher than second-order physical cumulants. According to the model the variance will become equal under the two measures only when $r_{t,T}$ follows the normal distribution ($k_{t,n}^P = 0$ for $n > 2$) and/or when the investor is risk-neutral ($\gamma = 0$). The empirically documented negative variance risk premium $k_{t,2}^P - k_{t,2}^Q < 0$ can be attributed by the model to negative skewness (i.e., $k_{t,3}^P < 0$) and/or positive excess kurtosis (i.e., $k_{t,4}^P > 0$). Equivalently, equation (9) indicates that the negative variance risk premium can be explained by the negative skewness (i.e., $k_{t,3}^Q < 0$) of the risk-neutral distribution (see Rompolis and Tzavalis (2010)).

Likewise, for $n = 3$, we get expressions for the third-order cumulant risk premium,

$$k_{t,3}^P - k_{t,3}^Q = \gamma k_{t,4}^P + \dots \quad (10)$$

and

$$k_{t,3}^P - k_{t,3}^Q = \gamma k_{t,4}^Q + \dots, \quad (11)$$

respectively, which is related to the difference of the skewness of the two distributions (also known in the literature as the skewness premium). Both formulas indicate that as soon as $k_{t,4}^Q > 0$ and $k_{t,4}^P > 0$ (i.e., the log-return distribution exhibits leptokurtosis under both measures) then the skewness premium is positive. By continuing this procedure for $n \geq 4$, we obtain expressions for risk premia corresponding to cumulants of increasingly higher order. As the set of these premia describe in essence the shape discrepancy between the two distributions, it becomes clear that the estimation of γ , practically depends on that information.

3 Econometric formulation

As already discussed, before we proceed to the estimation of the MRP, we first need to extract higher than first-order physical and risk-neutral cumulants and the PRRAC γ from our data. In this section we will describe the econometric formulation employed to that purpose.

3.1 Estimating higher-order physical cumulants

We estimate the 1-month forward-looking physical cumulants of the log-return distribution in a monthly frequency with the help of the Filtered Historical Simulation (FHS) as described in Barone-Adesi, Engle and Mancini (2008). This approach combines the well-known properties of a GARCH-family model (stochastic, mean-reverting and asymmetric volatility dynamics) with the empirical innovation distribution, which is able to fully capture various types of asymmetries and non-normalities observed in practice (see Wang, Zhang and Zhou (2015)). This way, we allow the log-return distribution to take on shapes that may not be possible under parametric specifications. It is no surprise, that methods of this type have been found in the recent literature to provide a

better basis for market-risk stress testing (see Basu (2011)). In essence, the method aims at the construction of the physical distribution of returns using efficiently the information contained in the historical path observed at a higher frequency relative to the horizon considered. It accomplishes that by combining inference from a calibrated conditional variance model and the empirical distribution of the innovations that drive the return process.

Let's denote as r_t the daily log-return for which we assume a general ARMA process ν_t , i.e.,

$$r_t = \nu_t + a_t \tag{12}$$

where $a_t = \sigma_t \cdot \varepsilon_t$ is conditionally heteroskedastic following mean-reverting, stochastic volatility dynamics and ε_t is a general white noise.

The first stage of the method is a straightforward, if rather cumbersome, econometric exercise of identifying the joint ARMA / variance model (ν_t / σ_t) that best explains the conditional mean and variance dynamics of the return process. Given that we aim to estimate a forward-looking physical distribution at each observation date, we allow ourselves to use only past data to infer the appropriate specification. We follow the Box-Jenkins procedure from scratch for each of these dates using a moving window that spans the 2 calendar years (about 500 observations $\{r_{t-499}, r_{t-498}, \dots, r_t\}$) that precede our data point. In order to arrive at the "best model" on each of these times, we examine a large number of candidate models that range from the simplest White Noise $r_t = \sigma \cdot \varepsilon_t$ to the more elaborate ARMA(p,q)-GARCH / EGARCH/ Threshold - GARCH (or GJR) specifications. All the latter address the volatility clustering effect of the observed returns, while Nelson's EGARCH (1991) and Glosten, Jagannathan and Runkle's Threshold - GARCH model (see Glosten, Jagannathan and Runkle (1993)) in addition allow for asymmetries in the return distribution. The candidate model set completed the GARCH-In-Mean specification that associates the conditional mean return with its conditional volatility. The total number of models considered on each data point is 688, half of which assume normal deviates in the log-likelihood function and the other half t -Student updates. From all the models that were deemed as "adequate"⁵ ones we choose the "best" based on the Bayesian Information Criterion (BIC) which generally tends to pick the most parsimonious specification.

Not surprisingly in almost 94% of our data points, the algorithm converged to the selection of an asymmetrical volatility model (EGARCH or GJR) while in 75% of the cases an ARMA(0,0) specification was chosen for the mean equation as shown in Table 1. Distribution-wise, the selected models are more evenly divided among normal (55%) and t -Student (45%) innovation distributions.

At the second stage of any FHS application, the fitted model serves as the filter that helps extract the primitive innovations ε_t that appeared to drive the process in its observed history. Therein exactly lies the strength of the method: The fact that the filtered innovations ε_t are not

⁵A model is characterized as "adequate" if its standardized residuals exhibit neither autocorrelation nor heteroskedasticity. We tested for the existence of the first using the Q-test and for the existence of the second using the ARCH-test.

bound to follow a specific form of randomness permits any singularity of the data that falls out of the standard norms of the parametric distribution to be modelled in the empirical density of the filtered innovations. As a result, the simulated distribution of $r_{t,T}$ that we generate at the third stage of the method is much richer in the values and patterns of cumulants that can be attained compared to its standard parametric counterparts.

The Monte Carlo simulation of that third stage is implemented by propagating random bootstraps from the set of the filtered innovations ε_t as sequences of primitive shocks that drive the fitted model from its current state at time t to its forward state at time $t + \tau$. The time-aggregation on each path i provided us with 100,000 simulated monthly returns, denoted as $r_{t,T}^{(i)}$, of which we estimate the cumulants of its distribution.⁶

The problem of choosing the length of the estimation window in our application is typical of the trade-off between having a sufficiently large sample to estimate a reasonably rich variance model and still retain only the most recent history so as to avoid producing "averaged" empirical distributions that bear little resemblance to the present. An important factor in our decision to use a moving window of about 500 daily returns rather than a longer or shorter one was the prevailing Risk-Management practice of using a history of about 2 years to calculate Value at Risk (e.g. RiskMetriks).

The main benefit of our flexible-model approach as opposed to using a single filter specification throughout the entire sample, as usually done in the literature (see Barone-Adesi, Engle and Mancini (2008), *inter alia*), is that we produce physical distributions, the cumulants of which, evolve in a much smoother way than otherwise. Indeed, if we are to calibrate the same model specification for all the months in our sample, then on those data points where this specification is smaller in size than the "locally optimal" mean/volatility model, the filtered innovations will have to be by construction autocorrelated or larger in magnitude so as to compensate for the omitted variables in the return dynamics. By allowing structural flexibility in the filter specification, we avoid both autocorrelation and outliers in the filtered innovations, which will in all certainty bias the inferred empirical distribution. On the other hand, in those cases where the single-model specification is larger in number of parameters than the "locally optimal" ones, the simulated variance of returns from that single-model will become inflated due to the redundant (statistically insignificant) model parameters that contaminate the sample with model noise.⁷

Note here that the FHS also generates an estimate of the mean of the physical distribution, i.e., $k_{t,1}^P$. However, this estimate is not economically meaningful, as it often takes negative values

⁶Starting from the non-central moment estimates $\mu_{t,n}^P = \frac{1}{N} \sum_{i=1}^N \left(r_{t,T}^{(i)} \right)^n$ of different order n , where $N = 100,000$, we calculate the n th-order cumulant of the physical distribution $k_{t,n}^P$ by applying the well-known k -Statistics estimators (see Kenney and Keeping (1951)).

⁷For example consider simulating an ARMA(1,1)/GARCH(1,2) in a specific subsample where the locally optimal model is only an ARMA(0,0)/GARCH(1,1) specification: The empirical distribution that this simulation will produce will be contaminated by a noise component, whose source are the statistically insignificant parameters of the larger model (ARMA terms and the second ARCH term), the value of which however small, will be in all likelihood different from zero.

being the sample average of the simulated monthly returns $r_{t,T}^{(i)}$. Nonetheless, this estimate, along with $\mu_{t,n}^P$ for $n \geq 2$, is used in the context of the k -Statistics estimators to calculate higher-order physical cumulants $k_{t,n}^P$ for $n \geq 2$. This poses no problem however, given that higher-order cumulants are mean-invariant. This implies that their values do not depend on the mean of the distribution, notwithstanding whether this mean is the ex-ante ER of the physical measure or the sample average of the simulated returns.

3.2 Retrieving risk-neutral cumulants from option prices

Ex-ante estimates of non-central risk-neutral moments $\mu_{t,n}^Q$ can be directly obtained from out-of-the-money (OTM) European call and put prices employing the formulas suggested by Bakshi, Kapadia and Madan (2003) for $n = 1, 2, 3, 4$ and extended by Rompolis and Tzavalis (2008) to any order n as

$$\begin{aligned} \mu_{t,1}^Q &= e^{(r_f - \delta)\tau} - 1 - e^{r_f\tau} \left[\int_{S_t}^{+\infty} \frac{1}{K^2} C_t(\tau, K) dK + \int_0^{S_t} \frac{1}{K^2} P_t(\tau, K) dK \right], \text{ for } n = 1, \\ \mu_{t,n}^Q &= e^{r_f\tau} \left\{ \int_{S_t}^{+\infty} \frac{n}{K^2} \left[\ln \left(\frac{K}{S_t} \right) \right]^{n-2} \left[n - 1 - \ln \left(\frac{K}{S_t} \right) \right] C_t(\tau, K) dK \right. \\ &\quad \left. + \int_0^{S_t} \frac{n}{K^2} \left[\ln \left(\frac{K}{S_t} \right) \right]^{n-2} \left[n - 1 - \ln \left(\frac{K}{S_t} \right) \right] P_t(\tau, K) dK \right\}, \text{ for } n \geq 2. \end{aligned} \quad (13)$$

where $C_t(\tau, K)$ and $P_t(\tau, K)$ denote the European call and put option prices with strike price K and maturity interval $\tau = T - t$. As before r_f denotes the annual return of the risk-free asset and δ denotes the continuously compounded dividend yield. As these formulas employ integrals of continuous functions to retrieve the values of the risk-neutral moments based on them, we can employ cubic splines to interpolate the implied by our option prices volatilities between two different points of the data. Due to the lack of option prices at 0 and $+\infty$, we can extrapolate the implied volatilities constantly over the intervals $(0, K_{\min}]$ and $[K_{\max}, +\infty)$, where K_{\min} and K_{\max} is the minimum and maximum strike prices given by our data, respectively. Ex-ante risk-neutral cumulants $k_{t,n}^Q$ can be obtained using the k -Statistics estimators in accordance with the non-central moment estimates given by (13).

3.3 Estimating the projected relative risk aversion coefficient

The next step of our procedure is to estimate the PRRAC. To this end, we can rely on the theoretical results of Section 2. More specifically, formula (3) or (4) for $n \geq 2$ can form a system of equations from which we can extract an estimate of γ . Following Bakshi and Madan (2006) we adopt a GMM estimation procedure. Let I_{t-1} be a set of instruments whose values are known at time $t - 1$. Then GMM estimation can be performed using the following orthogonality conditions implied by formula

(3):

$$E \left[\mathbf{k}_{t,N}^{P,Q} + \sum_{m=1}^M \mathbf{k}_{t,m+2,m+N}^P \frac{(-\gamma)^m}{m!} \middle| I_{t-1} \right] = \mathbf{0}, \quad (14)$$

where $\mathbf{k}_{t,N}^{P,Q} = (k_{t,2}^P - k_{t,2}^Q, \dots, k_{t,N}^P - k_{t,N}^Q)'$ and $\mathbf{k}_{t,m+2,m+N}^P = (k_{t,m+2}^P, \dots, k_{t,m+N}^P)'$. Similarly, the orthogonality conditions implied by (4) can be written as:

$$E \left[\mathbf{k}_{t,N}^{P,Q} - \sum_{m=1}^M \mathbf{k}_{t,m+2,m+N}^Q \frac{\gamma^m}{m!} \middle| I_{t-1} \right] = \mathbf{0}, \quad (15)$$

where $\mathbf{k}_{t,m+2,m+N}^Q = (k_{t,m+2}^Q, \dots, k_{t,m+N}^Q)'$.

Bakshi and Madan (2006) and Duan and Zhang (2014) used (14) for $N = 2$ and $M = 2$ (i.e., a second-order approximation of formula (8)) to estimate the PRRAC. For the same purpose, Rompolis and Tzavalis (2010) employed (15) for $N = 2$ and $M = 2$ (i.e., a second-order approximation of formula (9)). Both of these approaches provide a single-equation estimation of γ based on restricting the difference between the physical and risk-neutral variances by means of the physical or risk-neutral skewness and excess kurtosis..

We generalize this approach in two fronts. First, we increase the number of regressors M in (14) and (15). Theoretically, this will enable us to retrieve a consistent estimator of γ with a weaker set of instruments I_{t-1} . Second, we increase the number of equations N in (14) and (15). This step is very important for the purpose of our analysis, because, if additional information concerning the discrepancy between physical and risk-neutral distributions is important for the estimation of γ , then the inclusion of the additional equations will make the GMM estimator more efficient. This will happen simply because the additional moment restrictions imposed through these equations carry significant information for the shape of the pricing kernel. On the other hand, we could find that the equations corresponding to the differences in the higher-order cumulants between the two measures contain mostly noise which is independent of γ . In that case, the efficiency of the augmented system estimator would deteriorate relative to the single-equation benchmark, indicating that the PRRAC (according to the assumed pricing kernel) determines only the difference of the respective variances and not much else concerning the shape of the two distributions. If this is true, then the assumption of the existence of a monotonic pricing kernel would not be supported by the data.

4 Empirical analysis

In this section, using the aforementioned econometric formulation and data from the S&P 500 index for the period 1996-2010, we estimate the ex-ante MRP for a 1-month investment horizon.

4.1 The data

Our empirical analysis uses index daily returns and cross-section sets on European option prices for the S&P 500 index. We consider closing prices of put and call options for the third Wednesday of each month from January 1996 to October 2010. To address liquidity and quality issues in the prices of deep-OTM options we take a number of precautionary steps in the data pre-processing stage of our analysis. That is, to select our final sample, we apply several data filter. First, option quotes less than $3/8$ are excluded from the sample. These prices may not reflect true option value due to proximity to tick size. Second, we have excluded options contracts that had zero trading volume and/or open interest. Third, options violating the boundary conditions are eliminated from the sample. The purpose of these filters is to make our risk-neutral cumulants estimates as insensitive as possible to market microstructure and liquidity biases. The option data are downloaded from OptionMetrics Ivy DB. The maturity interval of these option prices is approximately equal to one calendar month (i.e. roughly twenty two trading days). Hence, the estimates of risk-neutral cumulants derived by the above option data sets concern distributions of returns that exhibit no overlap in time.

As the risk-free interest rate and dividend yield we use the estimates employed in the OptionMetrics calculations. The interest rate is derived from British Banker's Association LIBOR rates and settlement prices of Chicago Mercantile Exchange Eurodollar futures. The dividend yield is estimated by the put-call parity relation of at-the-money option contracts. We refer the reader to the Ivy DB reference manual for further details.

4.2 Physical and risk-neutral cumulants estimates

Table 2 presents summary statistics for the 1-month ahead physical and risk-neutral cumulants monthly estimates that we obtained in the first stage of our analysis. For expositional reasons we report only the descriptive statistics for the second (i.e., the variance), third and fourth-order cumulants. Note here, that in the estimation of γ we use physical and risk-neutral cumulants up to order 10. We also report summary statistics for the annualized volatility, the skewness and kurtosis coefficients of the log-return distribution under both measures. Figure 1 graphically presents the time series of the inferred physical and risk-neutral cumulants. The results in both the table and the figure indicate that the risk-neutral variance is generally higher than the physical one, which is consistent with previous findings in the literature. Third-order cumulant estimates, which account for the skewness of the distribution, are negative under both measures. However, $k_{t,3}^Q$ estimates are generally higher in absolute terms than $k_{t,3}^P$, suggesting that the risk-neutral distribution is more negatively skewed than the physical one. Fourth-order cumulant estimates are positive for both measures. The fact that $k_{t,4}^Q$ are generally higher than $k_{t,4}^P$ clearly attests to the relatively fatter tails of the risk-neutral distribution. The correlation coefficient between cumulants of the same order is very high and it decreases proportionally to n . Thus, any change in the shape of the physical

distribution should be also reflected, through a pricing kernel, to the shape of the risk-neutral one. An interesting feature in Figure 1 is the variability of risk-neutral and physical cumulants which consistently increase in absolute value during periods of financial crises (e.g. the Asian currency crisis, the Russian default, the burst of the Dot-Com bubble and Lehman Brother's default) or unexpected events (e.g. the attack on the World Trade Center on 9/11).

The average values of physical and risk-neutral cumulants estimates reported in Table 2 are in accordance with the predictions of the theoretical models (3) and (4). The observed negative average variance risk premium $k_{t,2}^P - k_{t,2}^Q$ can be explained by the negative average physical or risk-neutral third-order cumulant estimates and the positive physical or risk-neutral fourth-order cumulant estimates, according to equation (8) or (9), respectively. Similarly, the positive average third-order cumulant risk premium $k_{t,3}^P - k_{t,3}^Q$ can be explained by the positive average physical or risk-neutral fourth-order cumulant estimates, according to equation (10) or (11), respectively.

4.3 Projected relative risk aversion coefficient estimates

To estimate the PRRAC γ we employ orthogonality conditions (14) for $M = 2$ to 5 and $N = 2$ to 5. For $N = 2$ and $M = 2$, we estimate γ from the single-equation moment restriction

$$E \left[k_{t,2}^P - k_{t,2}^Q - \gamma k_{t,3}^P + \frac{\gamma^2}{2!} k_{t,4}^P \middle| I_{t-1} \right] = 0, \quad (16)$$

that explains variance risk premium $k_{t,2}^P - k_{t,2}^Q$ using only two additional higher than second-order physical cumulants, i.e. $k_{t,3}^P$ and $k_{t,4}^P$. Bakshi and Madan (2006) and Duan and Zhang (2014) adopted in effect this specification, which is why we consider this as our model comparison benchmark in the following sections. The competing estimation specifications are produced by sequentially increasing both the number of regressors and the number of equations in the system, setting $M = 3, 4$ and 5 and $N = 3, 4$ and 5, respectively. For example for $N = 2$ and $M = 5$ we estimate γ again from the single-equation orthogonality condition restricting the variance risk premium to include five additional higher than second-order physical cumulants, i.e.,

$$E \left[k_{t,2}^P - k_{t,2}^Q + \sum_{m=1}^5 \frac{(-\gamma)^m}{m!} k_{t,m+2}^P \middle| I_{t-1} \right] = 0. \quad (17)$$

For $N = 3$ and $M = 2$ we estimate γ by the following system of two equations given as

$$E \left[\begin{array}{l} k_{t,2}^P - k_{t,2}^Q - \gamma k_{t,3}^P + \frac{\gamma^2}{2!} k_{t,4}^P \\ k_{t,3}^P - k_{t,3}^Q - \gamma k_{t,4}^P + \frac{\gamma^2}{2!} k_{t,5}^P \end{array} \middle| I_{t-1} \right] = \mathbf{0}, \quad (18)$$

which now restricts the variance and third-order cumulant risk premia simultaneously. For robustness check, we also estimate γ for specifications of the same order, using this time orthogonality conditions (15). If the theoretical model is correctly specified, then the estimates of γ should be

qualitatively similar under both econometric setups (14) and (15).

Starting from the premise that the PRRAC is time-varying and related to business cycle indicators (see Bliss and Panigirtzoglou (2004)), we estimate γ using a 6-year moving window of data updated monthly. More specifically, the estimate of γ at the observation date t is obtained using 6-year data prior to and including t resulting in 72 monthly observations of the relevant variables for the GMM estimation on that data point.⁸ Note here that because of the 6-year rolling window of data used, the estimation period for γ spans the time period from December 2001 to October 2010. Three sets of instruments are used in the GMM estimation procedure for (14). The first contains the constant and one period lagged values of $k_{t,2}^Q$, $k_{t,3}^Q$, $k_{t,4}^Q$ and $k_{t,5}^Q$ for the first, second, third and fourth equation of the system, respectively. The second and third set of instruments use in addition to the first, two and two and three periods of lagged values of the same variables, respectively. For (15) we use a similar structure for the three sets of instruments where the risk-neutral cumulants are now replaced with the corresponding physical ones. The estimation results from the three sets of instruments are similar for both econometric specifications, therefore we only report those from using the third one.

Table 3 reports descriptive statistics of the estimation results, across the sample period, for both specifications, (14) and (15) and for different values of N and M . More precisely, it reports the average, standard deviation, minimum and maximum values of the estimates of γ . It also reports the average standard error and the average and minimum values of p -value of Hansen's overidentified restriction test. If the null hypothesis of this test is rejected then this will indicate that the estimated model is misspecified. The source of misspecification can be the existence of an alternative class of pricing kernels, the number of equations in the system, or the order of approximation of the theoretical formulas (3) and (4).

Several conclusions can be drawn from the results of Table 3. To start with, the estimates of γ , for all different specification examined, are intuitively sensible, statistically significant and comparable with the results of other recent studies that estimated γ jointly from stock and option data.⁹

Second, we observe that the inclusion of extra regressors (i.e. by increasing M) decreases the level of γ and its standard error. This is true for all N and M in (14) and for all N provided that $M \leq 4$ in (15). This result indicates that lower-order approximations of the theoretical formulas tend to produce inconsistent estimates of the PRRAC.¹⁰ The decrease in standard errors with

⁸We have also performed the GMM estimation procedure using different sizes of moving window of data. These preliminary results indicate that the 6-year moving window provides a smooth series of time-varying estimates of γ for both econometric specifications with the smallest number of time series observations. Even when we used a larger size moving window of data, the estimates of γ are similar to those reported in the paper.

⁹Several recent papers have estimated γ using stock and options data from the S&P 500 index. For example, Rosenberg and Engle (2002) reported an estimate of γ close to 7, Bliss and Panigirtzoglou (2004) reported an estimate close to 4, Rompolis and Tzavalis (2010) reported an estimate close to 1.3 and Duan and Zhang (2014) reported an estimate close to 4.

¹⁰A likely source of this inconsistency could be the fact that the instruments used are not orthogonal to the error term in low-order approximations. Indeed, the higher-order cumulants that are excluded from the regression, being

respect to M indicates that the additional regressors have explanatory power on the dependent variables. Only for specification (15) we observe that a fifth-order expansion ($M = 5$) of risk-neutral cumulants provide less accurate estimates of γ compared to lower-order approximations. This indicates that for specification (15) we cannot increase the order of expansion M arbitrarily. This is something to expect as risk-neutral cumulants estimates of a very high order calculated from option prices may be significantly biased.¹¹ In contrast, specification (14) does not seem to suffer from these constraints, as it is based on the expansion of physical cumulants. The latter correspond to the sample cumulants of the monthly returns simulated from the FHS method, and thus they are much better behaved. Thus, even if both theoretical models (3) and (4) provide an equivalent framework to estimate γ , formula (3), that this paper demonstrates, provide an econometric specification which tend to give more robust estimates of γ .

Third, the inclusion of new equations (i.e. increasing N) in the GMM estimation procedure further decreases standard errors. This is true for all M and for both specifications examined. As already argued, this result is very intuitive because it indicates that the inclusion of new information describing in detail the discrepancy between the physical and risk-neutral distribution makes the estimator of γ much more efficient.

Fourth, the overidentified restrictions afforded by the modeling structure are rejected for several observation dates at the 5% significance level for small numbers of N . However, for $N \geq 4$ (that is, for at least a three-equations system) these restrictions are not rejected across all observation dates. This evidence further suggest that the augmented model constitutes a correct specification of the data.

As an illustration, Figure 2 plots the series of the estimates of γ across the sample period for three different specifications. The first one corresponds to that used by Bakshi and Madan (2006) and Duan and Zhang (2014) (i.e., (14) for $N = 2$ and $M = 2$) which we use as the benchmark for the comparison we make. The second and third series of estimates γ are produced by higher-order specifications (i.e. setting $N = 4$ and $M = 4$ in (14) and (15), respectively).¹² Plot A of Figure 3 plots the estimates of the benchmark model against the 95% confidence interval produced by the more general specification (i.e., (14) for $N = 4$ and $M = 4$), while Plot B plots the estimates of the latter two series against the 95% confidence interval of the benchmark specification.

Inspection of these two figures confirms that the PRRAC varies smoothly across time, which is something already reported in the literature. The highest value for γ in our sample period is obtained on September 2007 for all three specifications. Not by coincidence, this observation

correlated with the lower-order ones and persistent, can be in effect correlated with the instruments. If the model includes a large number of cumulants, then the ones that enter the error term pose a relatively smaller problem given that they are less persistent, and thus they are more likely to be orthogonal to the instruments.

¹¹This is due to the fact that as the order of the estimated risk-neutral cumulant increases the dependence of its value on deep-OTM option prices which are not traded in the market also increases. In the numerical procedure that we use, these unobserved option prices are captured by the constant extrapolation scheme of the implied volatility function.

¹²Similar graphs are generated from the other estimated specifications of the model.

corresponds to the beginning of the recent financial crisis. It is in that turning point, when investors realized that a period of expansion has ended and the economy will move into recession, that they became more risk averse.¹³ From this period onward we observe a sharp decrease in the estimates of γ reaching its lowest value (i.e., close to 1.5) on October 2008, i.e., at the period following Lehman Brother’s default. This positive investors reaction could be due to the emergency government programs announced and implemented during September-October 2008 to assist the financial sector, including the temporary guarantee program of the US Treasury for money market funds, the Fed program to lend against high-quality asset-backed commercial papers, and most importantly the Troubled Asset Relief Program (TARP) and measures of quantitative easing such as large-scaled asset purchases (LSAPs). After this period, γ increases sharply close to 2.5 and remains constant until the end of the sample period.

Also in the same figures, we see all three series of the estimates of γ moving in conjunction. More notably, the series produced by the higher-order specifications (albeit from different theoretical model) being very close together appear lower in value, more stable and arguably more accurate as estimates of γ . In support of this last point in particular, the results reported in Plot A of Figure 3 show that the estimates of the benchmark specification are often located outside the (tighter) confidence interval of its higher-order counterpart; thus implying that Duan and Zhang’s (2014) approach leads to estimates of the PRRAC which strongly deviate from the true parameter, especially during the financial crisis period. In contrast, the estimates generated from the two more general specifications considered, lie by and large within the (wider) confidence interval of the benchmark (see Plot B of Figure 3).

Overall, the results described in this section signify that the accurate estimation of γ under the theoretical models (3) and (4) requires a higher-order approximation of the series expansions and a larger number of equations in the estimation procedure compared to the ones used by Bakshi and Madan (2006) and Duan and Zhang (2014). The fact that the estimates of γ are robust to the choice of the econometric specification (i.e, (14) and (15)) further supports this argument.

4.4 Ex-ante market risk premium estimates

Using the estimates of γ along with higher-order physical and risk-neutral cumulants we can compute ex-ante estimates of the S&P 500 index risk premium for each observation date for an investment horizon of 1 month. To do so, we employ a third-order approximation of formula (5) and (7).¹⁴ Table 4 reports sample descriptive statistics of annualized ex-ante MRP estimates for the different model specifications employed in the estimation of the PRRAC. Figure 4 plots the time

¹³In terms of the theoretical model this is due to the fact that the increase in the magnitude of risk-neutral cumulants on September 2007 could not be explained by an analogous increase of the physical ones.

¹⁴These approximations are given as:

$$k_{t,1}^P - k_{t,1}^Q \simeq \gamma k_{t,2}^P - \frac{\gamma^2}{2!} k_{t,3}^P + \frac{\gamma^3}{3!} k_{t,4}^P$$

series of these estimates for the three specifications used in Figure 2.

Several conclusions can be drawn from the results of Table 4 and Figure 4. First, the annualized MRP is always positive, time-varying and counter-cyclical, showing high variability, during the sample period, from around 1% to around 100%. As expected it increases during the two financial turmoil periods, i.e., the burst of the Dot-Com bubble in 2002 and the sub-prime mortgage crisis (2007 to 2009). During the later period we observe a steady increase of the MRP during the second semester of 2007. This is attributed to the high values of γ during this period and the steady increase of physical and risk-neutral cumulants. Its highest value, in our sample, is observed on November 2008 (around 100%). It is very interesting that the magnitude of this value is now dominated by the very high levels in the physical and risk-neutral cumulants, $k_{t,j}^P, k_{t,j}^Q$ for $j = 2, 3, 4$ (see Figure 1). This implies that, during the recent financial crisis changes in the MRP can be distilled into different set of factors: In 2007 MRP increased mainly because investors raised their level of aversion towards risk, while in 2008 it increased because investors perceived very significant probability of extreme losses. Second, the first-order autocorrelation, denoted as $R(-1)$, of the annualized MRP is highly positive indicating the persistence of it. Third, reflecting on the estimates of γ reported in Table 3, we observe that the values and variability of the MRP decreases as M and/or N increase. Also, the estimates of the MRP obtained by applying formulas (5) and (7) are close to each other, providing evidence that the theoretical model is consistent with the data.

The average, across our sample, annualized MRP which is around 14% for (14) and 17.5% for (15) can give us an estimate of the unconditional MRP. It worth comparing this estimate with that implied by the realized levels of higher-order physical cumulants. Using sample monthly returns from 2001 to 2010 we estimate a realized level of variance equal to 0.0022 and a realized level of third and fourth-order cumulants equal to -0.00008 and 0.000006, respectively. These estimates are significantly lower in magnitude compared to the average levels of the ex-ante physical cumulants reported in Table 2. For example, the realized third-order cumulant across the whole sample is 68% lower in absolute value compared to the average level of the conditional ex-ante third-order cumulant. Combining these sample statistics with a third-order approximation of equation (5), we obtain a value for the unconditional MRP close to 8%.¹⁵ This estimate approximately matches the historical average MRP of between 4% and 9% (depending on the sample period) reported by Mehra and Prescott (2003). Thus, the average MRP obtained from the ex-ante cumulant estimates is close to twice that obtained from the ex-post ones. This result provides evidence that a significant part of the ex-ante MRP is due to a level of risk which, even if it is priced in the market (in a

and

$$k_{t,1}^P - k_{t,1}^Q \simeq \gamma k_{t,2}^Q + \frac{\gamma^2}{2!} k_{t,3}^Q + \frac{\gamma^3}{3!} k_{t,4}^Q,$$

respectively. Note here that the inclusion of higher than fourth-order physical or risk-neutral cumulants in the above expansions has a marginal effect on the estimates of the MRP.

¹⁵This unconditional risk premium relies on an ex-post estimate of γ equal to 2.74. This is retrieved by employing the previous econometric framework, i.e, orthogonality conditions (14) for $M = 2$ and $N = 4$, for the full sample of physical and risk-neutral cumulants observations. This estimate of γ is robust to different econometric specifications of the model.

forward-looking basis), it accounts for bad states of the economy that will not be materialized in the sample (the well-known peso problem, see for example Brown, Goetzmann and Ross (1995) and Santa-Clara and Yan (2010), *inter alia*). This risk component is embedded in both option prices (which is reflected in the risk-neutral cumulants) and the shape of the forward-looking conditional distribution of market log-returns (which is summarized in the ex-ante physical cumulants) and as a quantity far exceeds the unconditional realized variation in stock market returns.

5 The estimated expected returns in the context of the present value identity

To examine the validity of the ER estimates we have obtained in the previous sections, we suggest a new approach based on the present value identity. This identity relates the observed dividend-price ratio to both ERs and expected dividend growth rates. The crucial benefit in undertaking this exercise draws from the fact that it produces results, which are by construction independent of the pricing kernel we have assumed. Equally importantly, it is applied on a dataset (i.e., the realized dividends of the S&P 500 index) that has not been used in the estimation of ERs, and as such it cannot have affected our estimates.

This part of our empirical analysis can also be seen as an alternative way of examining the relation between stock price movements, ERs and expected dividend growth rates. However, rather than rely on the traditional predictive regressions (see Cochrane (2011) for a review), we regress *current* estimates of market ERs, retrieved in the previous section, on the *current* dividend-price ratio controlling for the effect of dividend growth rates. This method is consistent with theory, which as a matter of fact is founded on a contemporaneous relation between dividend-price ratio and ER.

An important advantage of our approach is that, being essentially ex-ante, it is liable to capture the risk component related to the peso problem. On the contrary, a predictive regression based by construction on the ex-post sample, would have no way of doing this.

5.1 The present value identity

To avoid cluttering the reader with data, we present results for ER estimates produced by specifications (14) with $N = M = 2$ (referred to as model (M1) henceforth), and $N = M = 4$ (model (M2)) and (15) with $N = M = 4$ (model (M3)) for the remainder of the paper. Model (M1) is the benchmark model derived by Duan and Zhang (2014), while (M2) and (M3) should in theory be equivalent specifications in terms of their informational content.¹⁶

Let $r_{t,t+i} = \ln(S_{t+i}/S_t)$ denote the log-return without dividends between dates t and $t+i$, for $i = 1, 2, \dots$ months. Then, we define as $x_t \equiv k_{t,1}^P = E_t^P[r_{t,t+1}]$, i.e., the 1-month ahead ER estimated

¹⁶The results of this section are quite robust when derived from ER estimates obtained from system specifications of different order. These are available from the authors upon request.

at time t . Let also $\Delta d_{t,t+i} = \ln(D_{t+i}/D_t)$ be the realized dividend growth rate between dates t and $t+i$, $g_t = E_t^P [\Delta d_{t,t+1}]$ the 1-month ahead expected dividend growth rate and $dp_t = \ln(D_t/S_t)$ the log dividend-price ratio.

Starting from the identity $dp_t = r_{t,t+1} - \Delta d_{t,t+1} + dp_{t+1}$ and requiring that $\lim_{j \rightarrow \infty} dp_{t+j} = 0$, i.e., the dividend-price ratio does not explode, and finally take conditional expectations under physical measure P we can prove the following present value identity:

$$dp_t = \sum_{j=0}^{\infty} E_t^P [r_{t+j,t+j+1}] - \sum_{j=0}^{\infty} E_t^P [\Delta d_{t+j,t+j+1}]. \quad (19)$$

Formula (19) indicates that the dividend-price ratio is associated with both ERs and expected dividend growth rates. The dividend-price ratio increases (which means that price decreases) when ERs increase and/or expected dividend growth rates decrease. Note here that formula (19) is not exactly the same to the well-known Campbell-Shiller (1988) present value relation, as we use without-dividends returns instead of total returns. This is due to the fact that our model provides estimates of ERs without dividends as the terminal payoff of European options depend on the index level net of the accrued realized dividends. An additional advantage of this approach as compared to the traditional present value relation, is that formula (19) is an identity rather than a log-linear approximation.

By examining the properties (autocorrelation and partial autocorrelation functions) of the estimates of x_t given by models (M1), (M2) and (M3), we conclude that in all three cases x_t can be modeled as an $AR(1)$ process:¹⁷

$$x_{t+1} = \varphi_0 + \varphi_1 x_t + \varepsilon_{t+1}^x. \quad (20)$$

The least squares estimates of this model are given in Panel A of Table 5. Following the relevant literature we also assume that g_t follows an $AR(1)$ process (see Van Binsberger and Koijen (2010), *inter alia*), i.e.,

$$g_{t+1} = \theta_0 + \theta_1 g_t + \varepsilon_{t+1}^g. \quad (21)$$

Then, formula (19) can be written as:

$$dp_t = \kappa + \frac{1}{1 - \varphi_1} x_t - \frac{1}{1 - \theta_1} g_t, \quad (22)$$

where κ is a constant. The proof of the previous formula is provided in the Appendix. Formula (22) relates the dividend-price ratio with the 1-month ahead ER and dividend growth rate. As x_t is now an observed variable we can directly examine the effect of ERs and dividend growth rates

¹⁷This empirical result is also consistent with several studies which also assumed that $x_t \sim AR(1)$ process (see Cochrane (2008), *inter alia*).

on the dividend-price ratio by writing (22) as a contemporaneous linear regression model:

$$dp_t = \tilde{\kappa} + \frac{1}{1 - \varphi_1} x_t - \frac{\theta_1}{1 - \theta_1} \Delta d_{t-1,t} + u_t, \quad (23)$$

where $\tilde{\kappa}$ is a constant and $u_t = \frac{1}{1 - \theta_1} (\theta_1 \varepsilon_t^{\Delta d} - \varepsilon_t^g)$, where $\varepsilon_t^{\Delta d}$ is the forecast error of $\Delta d_{t-1,t}$.

5.2 Empirical application

To estimate equation (23) we use the three ER series, (M1), (M2) and (M3), described previously. We obtain monthly dividends for the S&P 500 index from Robert Shiller's website.¹⁸ Using these dividends and the index price at the day t for which x_t is estimated (i.e., the third Wednesday of each month) we calculate the dividend-price ratio. To estimate the econometric counterpart of formula (23) given as:

$$dp_t = \beta_0 + \beta_1 x_t + \beta_2 \Delta d_{t-1,t} + u_t, \quad (24)$$

we have to take into account the fact that at least one of the explanatory variables is, according to the theory laid out in the previous section, correlated with the error term u_t .¹⁹ Consequently, the OLS estimator should be biased and inconsistent. Moreover preliminary results, which can be provided by the authors upon request, indicate that both explanatory variables seem to be endogenous themselves implying that we cannot use simple OLS to determine their impact on the variation of the dividend-price ratio. We therefore use the weighted GMM estimator with the heteroskedasticity and autocorrelation consistent (HAC) weighting matrix sequentially updated until convergence is achieved.²⁰ Given that all variables are highly persistent we allow the error term u_t to follow an $AR(1)$ process. For robustness check, we also estimate (24) using a two-stage least squares (2-SLS) approach as it is widely known that the GMM estimator can exhibit a number of pathologies due to weak instrument identification problems.

The estimation results are reported in Panel B of Table 5. These indicate that both x_t and $\Delta d_{t-1,t}$ are statistically significant at the 1% level. Their variation can explain approximately 90% of the variation of dp_t . The coefficient estimates enter into the equation with the correct sign. More specifically, an increase in the 1-month ER is associated with an increase in the dividend-price ratio, whereas an increase in the dividend growth rate is followed by a decrease in the dp_t . More notably, the estimates of β_1 reported in Table 5 is very close to that expected by the $AR(1)$ structure of

¹⁸See <http://www.econ.yale.edu/~shiller/data.htm>.

¹⁹It is straightforward to verify that the contemporaneous model (23) posits that the covariance of error term u_t with $\Delta d_{t-1,t}$ is non-zero given that the former contains the forecast error of the latter.

²⁰The instruments used in the GMM estimation procedure are the 2 lags of all three variables, i.e. $(dp_{t-1}, dp_{t-2}, \Delta d_{t-2,t-1}, \Delta d_{t-3,t-2}, x_{t-1}, x_{t-2})$. This choice ensure that we are not missing information due to possible feedback effects between them and is in line with Fair's (1970) result. The latter indicates that when using the iterative Cochrane-Orcutt procedure to estimate a model, the lags of all dependent and independent variables have to be included in the instruments set so as to produce consistent estimates. The Eichenbaum, Hansen and Singleton (EHS) test which assesses the orthogonality condition of a subset of the instruments shows that the second lags of all variables are not necessary (the p -value of the null hypothesis was well above 0.10 in all three cases).

x_t , i.e., $\frac{1}{1-\varphi_1}$, for the ER series estimated by models (M2) and (M3). Indeed, for the case of (M2) and (M3) we have that $\frac{1}{1-\varphi_1}$ is equal to 3.7 and 4.76, respectively (see Panel A of Table 5). This is not what happens however, with the ER series produced by the benchmark model (M1) of Duan and Zhang (2014): While the level of persistence in fitted the $AR(1)$ model calls for an expected coefficient of 3.57 the estimated β_1 (equal to 2.26) is almost 2 standard errors away from that value, implying that this particular series fails to satisfy the present value identity. Apart from this, our results generally indicate that, at least for our sample period, the variation of the dividend-price ratio is due to the variation of both ERs and expected dividend growth rates.

6 Conclusion

This paper finds that the estimation of the market ER and the associated ex-ante MRP benefits significantly from the inclusion of additional information about the shape of physical and risk-neutral distributions, and the shape discrepancy between the two.

We can describe the benefits along two dimensions: First, statistically, GMM specification tests of overidentifying restrictions indicate that a single equation framework for the estimation of the PRRAC, as used in the previous literature, is not always supported by the data. When, on the contrary, we exploit additional information by increasing the number of equations in the system, the GMM estimator becomes considerably more efficient. At the same time, model specifications of low-order approximation, that lack in the shape information embedded in higher-order cumulants, appear to produce inconsistent estimates of the PRRAC in certain cases.

Second, economically, our empirical evidence suggests that ER estimates produced from the augmented information sets, are the more plausible, given that they satisfy, almost perfectly, the present value identity. It is important to stress here, that this approach can be considered as an objective measure of comparison as it is independent of the framework employed for the estimation of ERs.

Our results are very important because they imply that a richer description of the shape of return distributions improves both statistically and economically the estimation of the ER and the associated MRP. Given the central role of the MRP in asset pricing theory and risk-return decision making, the suggested estimation method can be of particular interest both for researchers and practitioners.

A Appendix

In this appendix, we provide proofs of the main theoretical results of the paper

A.1 Proof of Proposition 1

To prove formula (3) take the n^{th} -order derivative of (2) with respect to u evaluated at $u = 0$ which gives:

$$\left[\frac{d^n m_t^Q(u)}{du^n} \right]_{u=0} = \left[\frac{d^n m_t^P(u - \gamma)}{du^n} \right]_{u=0}, \quad \forall n \in \mathbb{N}. \quad (25)$$

By definition the n^{th} -order cumulant of $r_{t,T}$ under the risk-neutral measure is equal to $k_{t,n}^Q = \left[\frac{d^n m_t^Q(u)}{du^n} \right]_{u=0}$. Also writing $m_t^P(u - \gamma)$ in a power series expansion as

$$m_t^P(u - \gamma) = \sum_{m=1}^{\infty} k_{t,m}^P \frac{(u - \gamma)^m}{m!} \quad (26)$$

and calculating the n^{th} -order derivative yields

$$\left[\frac{d^n m_t^P(u - \gamma)}{du^n} \right]_{u=0} = \sum_{m=0}^{\infty} k_{t,n+m}^P \frac{(-\gamma)^m}{m!}. \quad (27)$$

Substituting formulas (26) and (27) into (25) yields

$$k_{t,n}^Q = \sum_{m=0}^{\infty} k_{t,n+m}^P \frac{(-\gamma)^m}{m!}, \quad \forall n \in \mathbb{N},$$

which is exactly formula (3).

A.2 Proof of formula (22)

Given that x_t and g_t follow an $AR(1)$ process we can easily prove that:

$$E_t^P [r_{t+j,t+j+1}] = E_t^P [x_{t+j}] = \varphi_0 \frac{1 - \varphi_1^j}{1 - \varphi_1} + \varphi_1^j x_t, \quad (28)$$

and

$$E_t^P [\Delta d_{t+j,t+j+1}] = E_t^P [g_{t+j}] = \theta_0 \frac{1 - \theta_1^j}{1 - \theta_1} + \theta_1^j g_t. \quad (29)$$

Also note that if we take conditional expectations under P on the identity $dp_t = r_{t,t+1} - \Delta d_{t,t+1} + dp_{t+1}$ we have that $dp_t = x_t - g_t + E_t^P [dp_{t+1}]$. Taking unconditional expectations yields $E^P [dp_t] = E^P [x_t] - E^P [g_t] + E^P [dp_{t+1}]$, which is equivalent to $E^P [x_t] = E^P [g_t] \Rightarrow \frac{\varphi_0}{1 - \varphi_1} = \frac{\theta_0}{1 - \theta_1}$, if the dividend-price ratio is a stationary process.

Substituting formulas (28) and (29) and the stationarity condition for dp_t into formula (19) yields:

$$dp_t = \kappa + \frac{1}{1 - \varphi_1} x_t - \frac{1}{1 - \theta_1} g_t,$$

where $\kappa = -\frac{\varphi_0}{(1-\varphi_1)^2} + \frac{\theta_0}{(1-\theta_1)^2}$, which is exactly formula (22).

References

- [1] Ait-Sahalia, Y., and A. W. Lo, 2000, Nonparametric risk management and implied risk aversion, *Journal of Econometrics* 94, 9-51.
- [2] 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, 101-143.
- [3] Bakshi, G., and D. Madan, 2006, A theory of volatility spread, *Management Science* 52, 1945-1956.
- [4] Barone-Adesi, G., R. F. Engle, and L. Mancini, 2008, A GARCH option pricing model with filtered historical simulation, *Review of Financial Studies* 21, 1223-1258.
- [5] Basu, S., 2011, Comparing simulation models for market risk stress testing, *European Journal of Operational Research* 213, 329-339.
- [6] Binsberger, Van, J. H., and R. S. Koijen, 2010, Predictive regressions: A present-value approach, *Journal of Finance* 65, 1439-1471.
- [7] Bliss, R., and N. Panigirtzoglou, 2004, Option-implied risk aversion estimates, *Journal of Finance* 59, 407-446.
- [8] Brown, S. J., W. N. Goetzmann, and S. A. Ross, 1995, Survival, *Journal of Finance* 50, 853-873.
- [9] Campbell, J., and R. Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195-228.
- [10] Chalamandaris, G., and L. S. Rompolis, 2012, Exploring the role of the realized return distribution in the formation of the implied volatility smile, *Journal of Banking and Finance* 36, 1028-1044.
- [11] Chen, L., Z. Da, and X. Zhao, 2013, What drives stock price movements? *Review of Financial Studies*, 26, 841-876.
- [12] Christensen, B. J., and N.R. Prabhala, 1998, The relation between implied and realized volatility, *Journal of Financial Economics* 50, 125-150.
- [13] Cochrane, J. H., 2008, The dog that did not bark: A defense of return predictability, *Review of Financial Studies* 21, 1533-1575.
- [14] Cochrane, J. H., 2011, Discount rates, *Journal of Finance* 66, 1047-1108.
- [15] Duan, J. C., and W. Zhang, 2014, Forward-looking market risk premium, *Management Science* 60, 521-538.

- [16] Fair, R. C., 1970, The estimation of simultaneous equation models with lagged endogenous variables and first order serially correlated errors, *Econometrica* 38, 507-516.
- [17] Glosten, L. R., R. Jagannathan, and D. E. Runkle, 1993, On the relation between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance* 48, 1779-1801.
- [18] Graham, J. R., and C. R. Harvey, 2007, The equity risk premium in January 2007: Evidence from the global CFO outlook survey, *The ICFAI Journal of Financial Risk Management* 4, 46-61.
- [19] Goh, J., K. G. Lim, M. Sim, and W. Zhang, 2012, Portfolio value-at-risk optimization for asymmetrically distributed asset returns, *European Journal of Operational Research* 221, 397-406.
- [20] Kenney, J. F., and E. S. Keeping, 1951, *Mathematics of Statistics Pt. 2.*, D. Van Nostrand, New York.
- [21] Koijen, R., and S. Van Nieuwerburgh, 2011, Predictability of returns and cash flows, *Annual Review of Financial Economics* 3, 467-491.
- [22] Mehra, R., and E. C. Prescott, 2003, The equity premium in retrospect, in G. M. Constantinides, M. Harris, and R. M. Stulz (Eds.), *Handbook of the Economics of Finance*, Elsevier, Amsterdam.
- [23] Merton, R. C., 1980, On estimating the expected return on the market: An exploratory investigation, *Journal of Financial Economics* 8, 323-361.
- [24] Nelson, D. B., 1991, Conditional heteroskedasticity in asset returns: A new approach, *Econometrica* 59, 347-370.
- [25] Pastor, L., M. Sinha, and B. Swaminathan, 2008, Estimating the intertemporal risk-return tradeoff using the implied cost of capital, *Journal of Finance* 63, 2859-2897.
- [26] Rompolis, L. S., Tzavalis, E., 2008, Recovering risk neutral densities from option prices: A new approach, *Journal of Financial and Quantitative Analysis* 43, 1037-1054.
- [27] Rompolis, L. S., and E. Tzavalis, 2010, Risk premium effects on implied volatility regressions, *Journal of Financial Research* 33, 125-151.
- [28] Rosenberg, J., and R. Engle, 2002, Empirical pricing kernels, *Journal of Financial Economics*, 64, 341-371.
- [29] Santa-Clara, P., and S. Yan, 2010, Crashes, volatility and the equity premium: Lessons from S&P 500 options, *The Review of Economics and Statistics* 92, 435-451.

- [30] Wang, P., B. Zhang, and Y. Zhou, 2015, Asymmetries in stock markets, *European Journal of Operational Research* 241, 749-762.
- [31] Welch, I., 2000, Views of financial economists on the equity premium and on professional controversies, *Journal of Business* 73, 501-537.

Table 1: Selected model frequency

| Mean Equation | Frequency | Variance Equation | Frequency |
|---------------|-----------|-------------------|-----------|
| ARMA(0,0) | 75% | EGARCH(1,1) | 70% |
| ARMA(0,1) | 14% | EGARCH(2,2) | 15% |
| ARMA(1,0) | 8% | GJR(1,1) | 9% |
| ARMA(1,1) | 3% | GARCH(1,1) | 6% |

This table reports the specifications selected (among the 688 candidate models on each observation date) during our flexible-model Filtered Historical Simulation procedure, as well as their corresponding frequency of selection. The first two columns display the models selected for the mean equation and the last two the models selected for the variance equation.

Table 2: Descriptive statistics

| | Risk-neutral | Physical | Correlation (%) |
|-----------------------|-------------------------|------------------------|-----------------|
| Variance | 0.004638 (0.005310) | 0.003112 (0.003583) | 94 |
| 3rd-order cumulant | -0.000605 (0.001226) | -0.00027 (0.000588) | 93 |
| 4th-order cumulant | 0.000164 (0.000398) | 0.000068 (0.000223) | 87 |
| Annualized Volatility | 21.68 | 17.61 | |
| Skewness | -1.53 | -0.96 | |
| Kurtosis | 7.99 | 5.37 | |

This table reports average values of second, third and fourth-order cumulants of the physical and risk-neutral distribution. Standard deviations are in parentheses. The last column of the table presents the correlation coefficient between the respective risk-neutral and physical cumulants. The table also presents the average annualized volatility (in percentage points) and the average skewness and kurtosis coefficients derived from the physical and risk-neutral cumulants, respectively. Risk-neutral cumulants are estimated directly from S&P 500 index option prices with approximately 1 month time-to-maturity using equation (13). Physical cumulants are estimated by the FHS method employed to daily index returns. The estimation period span from January 1996 to October 2010.

Table 3: Projected relative risk aversion coefficient estimates

| N | 2 | | | | | 3 | | | | | 4 | | | | | 5 | | | | |
|---------------------|---------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|---|---|--|
| | 2 | 3 | 4 | 5 | | 2 | 3 | 4 | 5 | | 2 | 3 | 4 | 5 | | 2 | 3 | 4 | 5 | |
| Panel A: Model (14) | | | | | | | | | | | | | | | | | | | | |
| γ | Avg. | 4.33 | 3.91 | 3.72 | 3.62 | 4.51 | 3.99 | 3.73 | 3.55 | 4.46 | 4.01 | 3.71 | 3.48 | 4.41 | 3.83 | 3.46 | 3.21 | | | |
| | St.Dev. | 1.28 | 1.07 | 0.98 | 0.92 | 1.26 | 1.09 | 1.00 | 0.92 | 1.25 | 1.14 | 1.04 | 0.94 | 1.23 | 1.08 | 0.95 | 0.86 | | | |
| SE | Min | 1.99 | 1.83 | 1.80 | 1.79 | 1.69 | 1.60 | 1.58 | 1.58 | 1.69 | 1.60 | 1.58 | 1.57 | 1.62 | 1.53 | 1.52 | 1.50 | | | |
| | Max | 8.25 | 7.05 | 6.68 | 6.52 | 7.40 | 6.42 | 5.97 | 5.77 | 6.76 | 6.15 | 5.80 | 5.60 | 6.76 | 5.89 | 5.58 | 5.45 | | | |
| p -value | Avg. | 0.66 | 0.53 | 0.47 | 0.43 | 0.53 | 0.42 | 0.37 | 0.35 | 0.46 | 0.33 | 0.26 | 0.24 | 0.28 | 0.21 | 0.18 | 0.17 | | | |
| | Min | 0.46 | 0.46 | 0.45 | 0.45 | 0.36 | 0.34 | 0.34 | 0.33 | 0.41 | 0.42 | 0.41 | 0.40 | 0.46 | 0.47 | 0.47 | 0.46 | | | |
| | Min | 0.01 | 0.01 | 0.01 | 0.01 | 0.04 | 0.03 | 0.03 | 0.03 | 0.13 | 0.13 | 0.13 | 0.13 | 0.28 | 0.27 | 0.27 | 0.27 | | | |
| Panel B: Model (15) | | | | | | | | | | | | | | | | | | | | |
| γ | Avg. | 4.28 | 4.01 | 4.49 | 5.10 | 3.71 | 3.70 | 3.87 | 4.23 | 3.94 | 3.62 | 3.98 | 4.06 | 3.81 | 3.73 | 4.00 | 4.25 | | | |
| | St.Dev. | 0.84 | 0.97 | 1.27 | 1.78 | 1.52 | 0.77 | 0.86 | 1.46 | 1.52 | 0.71 | 0.88 | 1.19 | 1.69 | 0.76 | 0.88 | 1.18 | | | |
| SE | Min | 2.84 | 2.36 | 2.42 | 2.51 | 1.81 | 2.34 | 2.46 | 2.53 | 1.96 | 2.37 | 2.46 | 2.53 | 1.42 | 2.37 | 2.44 | 2.53 | | | |
| | Max | 6.34 | 6.39 | 7.04 | 8.38 | 9.83 | 5.78 | 5.57 | 7.71 | 9.10 | 5.47 | 5.57 | 7.55 | 9.97 | 5.51 | 5.41 | 7.62 | | | |
| p -value | Avg. | 1.81 | 0.50 | 0.69 | 0.73 | 0.74 | 0.25 | 0.23 | 0.33 | 0.26 | 0.18 | 0.11 | 0.29 | 0.08 | 0.05 | 0.04 | 0.09 | | | |
| | Min | 0.20 | 0.32 | 0.50 | 0.59 | 0.35 | 0.23 | 0.35 | 0.34 | 0.46 | 0.22 | 0.33 | 0.34 | 0.42 | 0.36 | 0.34 | 0.42 | | | |
| | Min | 0.01 | 0.01 | 0.01 | 0.01 | 0.03 | 0.04 | 0.03 | 0.03 | 0.11 | 0.10 | 0.09 | 0.10 | 0.26 | 0.29 | 0.29 | 0.29 | | | |

This table reports descriptive statistics of the GMM estimation results for both specifications, (14) and (15), and for different values of N and M . More precisely, it reports the average, standard deviation, minimum and maximum values of the estimates of γ . It also reports the average standard error (denoted as SE) and the average and minimum values of p -value of Hansen's overidentified restriction test. For each observation date we use a 6-year moving window of data to estimate γ . The instruments used with orthogonality conditions (14) are the constant and one, two and three periods lagged values of $k_{t,2}^Q$, $k_{t,3}^Q$, $k_{t,4}^Q$ and $k_{t,5}^Q$ for the first, second, third and fourth equation of the system, respectively. The instruments used with orthogonality conditions (15) are the constant and one, two and three periods lagged values of $k_{t,2}^P$, $k_{t,3}^P$, $k_{t,4}^P$ and $k_{t,5}^P$ for the first, second, third and fourth equation of the system, respectively. Standard errors are corrected for heteroskedasticity and serial correlation. The estimation period span from December 2001 to October 2010.

Table 4: Annualized ex-ante market risk premium estimates (in %)

| N | 2 | | | | | 3 | | | | | 4 | | | | | 5 | | | | |
|---------|----------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|---|---|---|--|
| | 2 | 3 | 4 | 5 | | 2 | 3 | 4 | 5 | | 2 | 3 | 4 | 5 | | 2 | 3 | 4 | 5 | |
| M | | | | | | | | | | | | | | | | | | | | |
| Avg. | 17.00 | 15.08 | 14.37 | 14.01 | 17.29 | 14.85 | 13.78 | 13.22 | 16.53 | 14.42 | 13.37 | 12.68 | 16.10 | 13.82 | 12.59 | 11.84 | | | | |
| St.Dev. | 18.65 | 16.43 | 15.75 | 15.48 | 18.14 | 15.50 | 14.60 | 14.27 | 16.98 | 14.77 | 14.05 | 13.69 | 16.26 | 14.29 | 13.52 | 13.17 | | | | |
| Min | 1.50 | 1.38 | 1.32 | 1.29 | 1.58 | 1.42 | 1.34 | 1.28 | 1.64 | 1.51 | 1.40 | 1.31 | 1.71 | 1.49 | 1.35 | 1.24 | | | | |
| Max | 113.0 | 100.7 | 97.29 | 96.23 | 108.8 | 96.50 | 93.18 | 92.26 | 106.6 | 94.30 | 91.04 | 90.23 | 103.7 | 91.84 | 88.89 | 88.32 | | | | |
| $R(-1)$ | 74.47 | 75.22 | 75.42 | 75.46 | 73.86 | 74.81 | 75.22 | 75.40 | 74.01 | 75.30 | 75.73 | 75.80 | 75.12 | 76.02 | 76.20 | 76.14 | | | | |
| | Panel A: Model (14), Formula (5) | | | | | | | | | | | | | | | | | | | |
| | Panel B: Model (15), Formula (7) | | | | | | | | | | | | | | | | | | | |
| Avg. | 19.55 | 18.53 | 18.51 | 21.16 | 16.05 | 16.90 | 16.79 | 17.93 | 16.92 | 16.37 | 16.99 | 17.53 | 16.79 | 16.51 | 17.04 | 17.91 | | | | |
| St.Dev. | 25.79 | 25.78 | 19.48 | 26.85 | 17.66 | 19.85 | 16.94 | 17.92 | 17.46 | 19.02 | 16.80 | 17.66 | 17.38 | 18.46 | 16.85 | 17.53 | | | | |
| Min | 3.88 | 3.15 | 3.72 | 4.05 | 1.91 | 2.88 | 3.14 | 2.95 | 2.28 | 2.93 | 3.23 | 3.21 | 1.87 | 3.08 | 3.31 | 3.39 | | | | |
| Max | 184.6 | 208.4 | 139.7 | 197.7 | 114.8 | 138.6 | 113.1 | 118.9 | 110.8 | 132.9 | 114.5 | 119.1 | 109.6 | 130.7 | 114.3 | 118.6 | | | | |
| $R(-1)$ | 77.34 | 71.47 | 78.41 | 72.10 | 79.71 | 79.16 | 81.22 | 80.43 | 70.62 | 78.83 | 80.33 | 80.64 | 75.62 | 79.04 | 80.32 | 80.20 | | | | |

This table reports sample descriptive statistics of the annualized ex-ante market risk premium estimates (in percentage points) of the S&P 500 index for the different model specifications used for the estimation of the projected relative risk aversion coefficient. It reports the average, standard deviation, minimum and maximum values of these estimates and their first-order autocorrelation, denoted as $R(-1)$. Panel A reports estimates of the ex-ante market risk premium based on estimates of γ retrieved from specification (14) and the third-order approximation of formula (5), while Panel B reports estimates of it based on a third-order approximation of formula (7). In this latter case the GMM estimates of γ are based on orthogonality conditions (15). The estimation period span from December 2001 to October 2010.

Table 5: Estimation results for the present value identity

| Model | M1 | M2 | M3 | | | |
|---|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| Panel A: LS estimates of model (20) | | | | | | |
| φ_0 | 0.0037 (3.08) | 0.0028 (3.11) | 0.0027 (2.7) | | | |
| φ_1 | 0.72 (10.58) | 0.73 (10.89) | 0.79 (12.95) | | | |
| R^2 | 0.52 | 0.53 | 0.62 | | | |
| $\frac{1}{1-\varphi_1}$ | 3.57 | 3.7 | 4.76 | | | |
| Panel B: Estimation results for equation (24) | | | | | | |
| | W-GMM | 2-SLS | W-GMM | 2-SLS | W-GMM | 2-SLS |
| constant | -3.89 (-76.67) | -3.92 (-73.10) | -3.86 (-84.49) | -3.92 (-72.09) | -3.94 (-72.71) | -3.94 (-83.63) |
| x_t | 2.26 (3.61) | 2.51 (3.71) | 3.16 (2.97) | 3.73 (3.97) | 4.48 (6.18) | 4.54 (5.59) |
| $\Delta d_{t-1,t}$ | -18.91 (-3.06) | -14.59 (-2.23) | -21.37 (-3.59) | -14.66 (-2.18) | -15.80 (-2.26) | -15.12 (-3.04) |
| ρ | 0.85 (18.17) | 0.89 (13.77) | 0.85 (14.74) | 0.89 (13.64) | 0.85 (16.74) | 0.88 (16.74) |
| R^2 | 0.88 | 0.91 | 0.86 | 0.91 | 0.92 | 0.93 |
| DW | 1.77 | 1.74 | 1.78 | 1.74 | 1.87 | 1.92 |

This table reports the estimation results for the present value identity. Panel A reports least squares (LS) estimates of model (20). Panel B reports the results of estimating equation (24). The ERs used in this empirical application were provided by three different system specifications. The first one (denoted as (M1)) corresponds to specification (14) with $N = 2, M = 2$. The second one (denoted as (M2)) corresponds to the same specification with $N = 4, M = 4$, and the third one (denoted as (M3)) corresponds to specification (15) of the same size. We estimate equation (24) using both the weighted GMM estimator (W-GMM) and the two-stage least squares estimator (2-SLS) allowing the error term to follow an $AR(1)$ process with autoregressive coefficient ρ . In both cases, the instrument list contains the two lags of the dependent and all independent variables. In both panels t -statistics are reported in parentheses. DW denotes the Durbin-Watson test statistic. The estimation period span from December 2001 to October 2010.

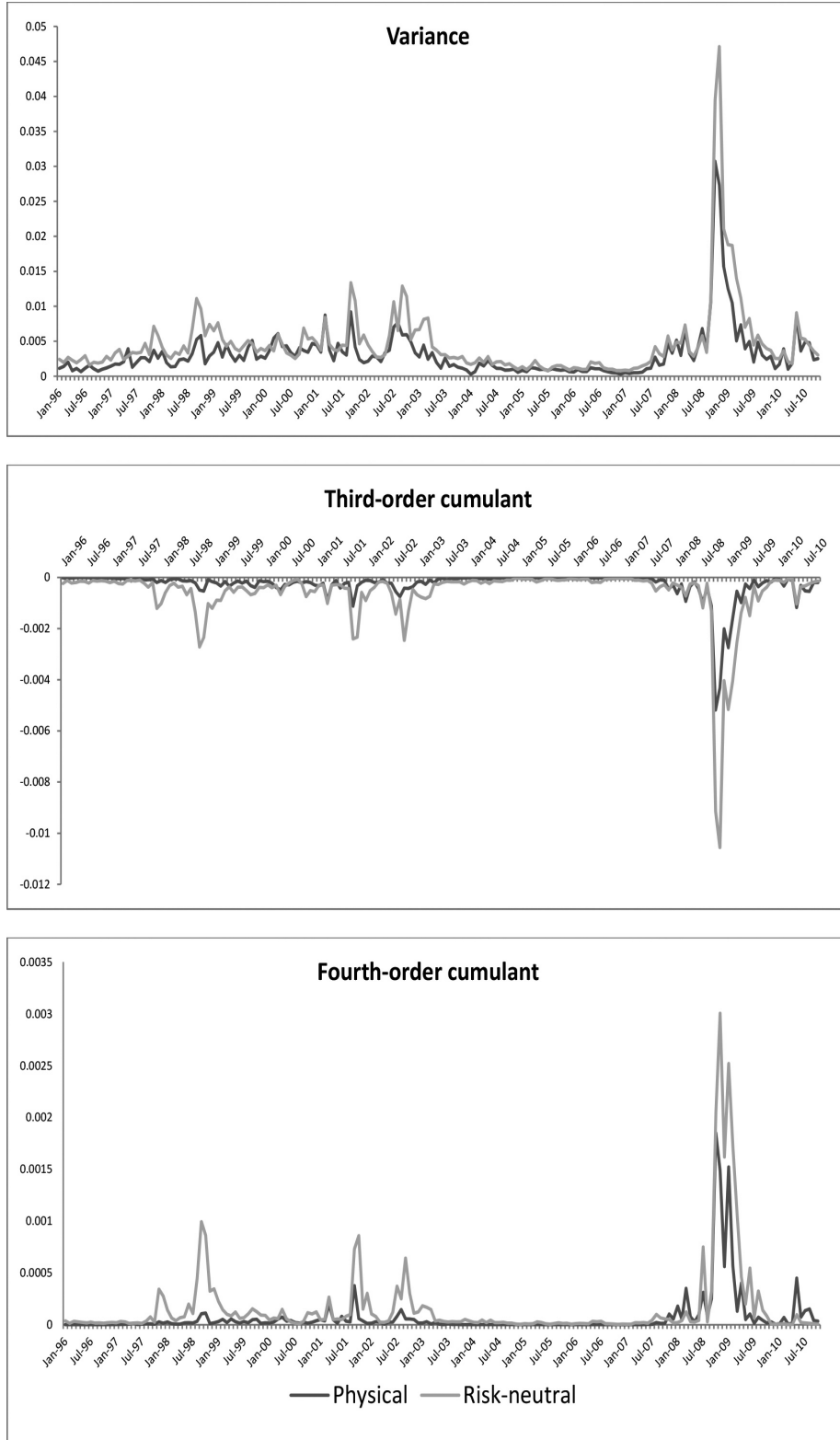


Figure 1: Time series monthly estimates of conditional variance, third and fourth-order cumulants of log-return distribution under physical and risk-neutral distributions from January 1996 to October 2010.

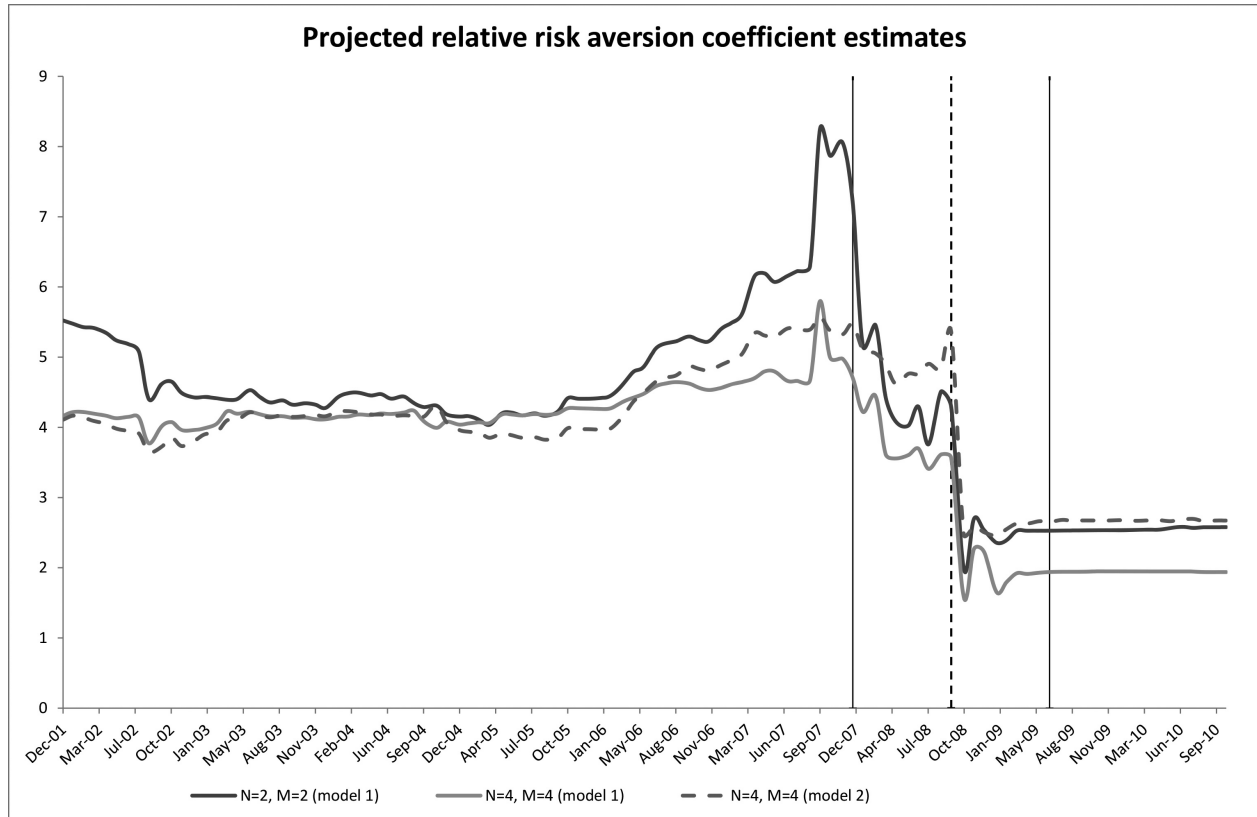


Figure 2: Projected relative risk aversion coefficient estimates, December 2001 to October 2010. Model 1 corresponds to econometric specification (14), while model 2 corresponds to (15). The vertical solid lines indicate the recent NBER recession period. The vertical dashed line indicate Lehman Brother’s default date, i.e, September 2008.

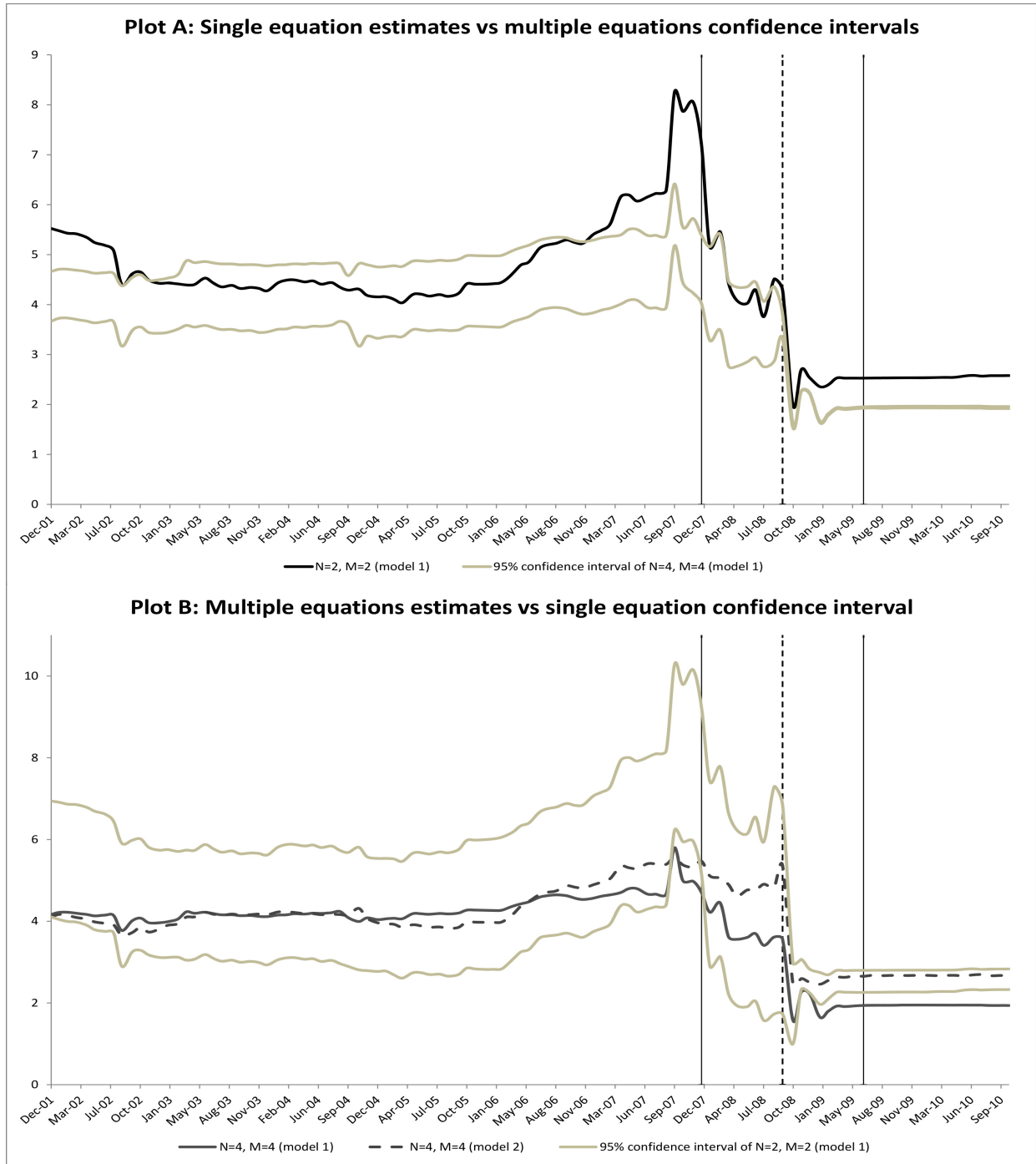


Figure 3: Projected relative risk aversion estimates and confidence intervals, December 2001 to October 2010. Plot A presents the estimates of the benchmark model (i.e., (14) with $N = M = 2$) against the 95% confidence interval produced by its higher-order counterpart with $N = M = 4$. Plot B presents the estimates of specifications (15) and (14) with $N = M = 4$ against the 95% confidence interval of the benchmark model. The vertical solid lines indicate the recent NBER recession period. The vertical dashed line indicate Lehman Brother's default date, i.e, September 2008.

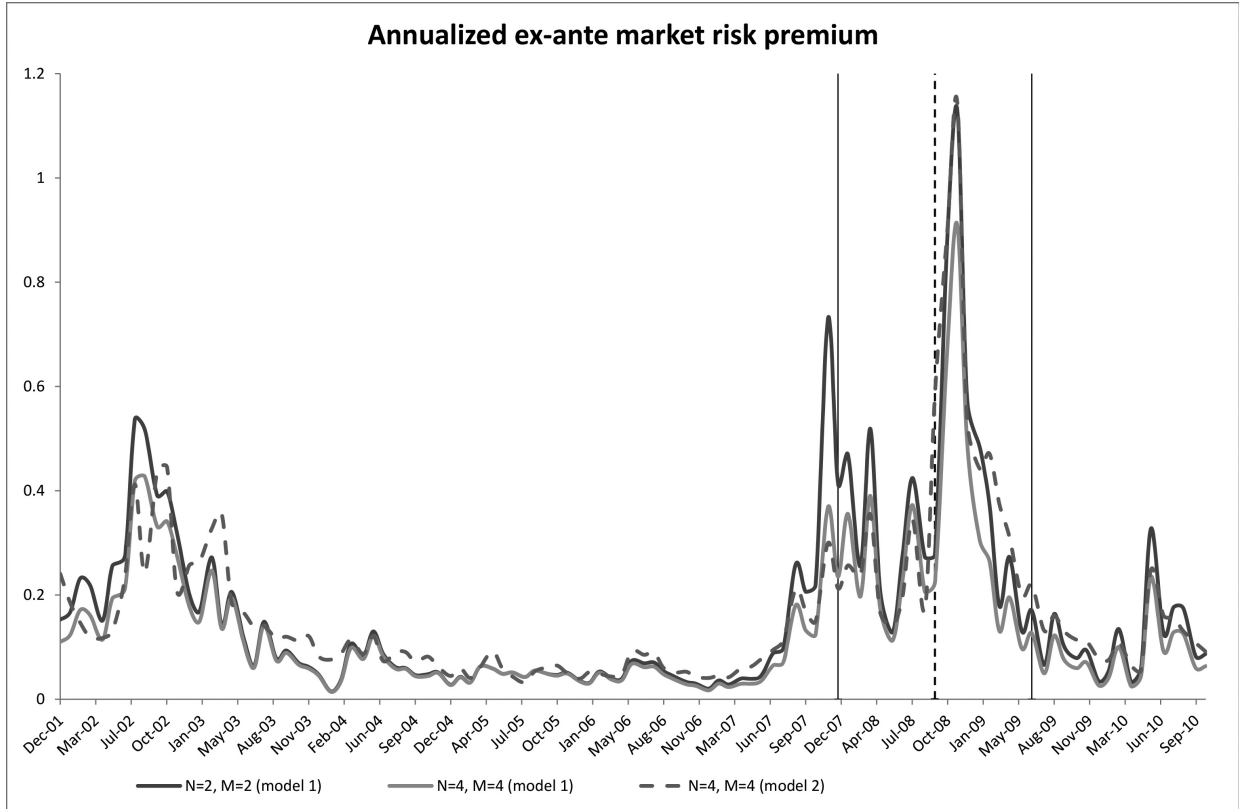


Figure 4: Annualized ex-ante risk premium estimates of the S&P 500 index produced by different specifications, December 2001 to October 2010. Model 1 corresponds to econometric specification (14), while model 2 corresponds to (15). The vertical solid lines indicate the recent NBER recession period. The vertical dashed line indicate Lehman Brother’s default date, i.e, September 2008.