Disentangling the Effects of Heterogeneous Beliefs and Preferences on Asset Prices*

Fousseni Chabi-Yo[†] Eric Ghysels[‡] Eric Renault[§]

First draft: February 2006,
This version: November 7, 2006

Abstract

Several authors have recognized that pricing kernels should reward assets for their contribution to the higher order moments, such as skewness and kurtosis, of the return on the market portfolio in order to successfully explain individual asset risk premiums. The main drawback is that the resulting empirical pricing kernels typically do not map directly into a fundamental preference-based asset pricing model, unlike pricing kernels that are based on CRRA or other utility functions. It is the purpose of this paper to close this gap. We go beyond the representative agent paradigm and provide a general framework that maps heterogeneity of beliefs and preferences into polynomial expansions of the pricing kernel and obtain a structural interpretation of the empirical pricing kernels involving skewness and kurtosis. Moreover, the structural interpretation of the pricing kernel allows us to disentangle the effects of heterogeneous beliefs and preferences on asset prices. To handle the various sources of heterogeneity, beliefs and preferences that is, we follow the framework of Samuelson (1970) and its recent generalization of Chabi-Yo, Leisen, and Renault (2006). This generic approach allows us to derive, for risks that are infinitely small, optimal shares of wealth invested in each security that coincide with those of a Mean-Variance-Skewness-Kurtosis optimizing agent. Through these local approximations we are able to tease out the various sources of risk. A by-product of our analysis is an approximate aggregation theory for general equilibrium heterogeneous agent asset pricing models.

*We thank seminar participants at CREST and UBC for helpful comments. Address for correspondence: Eric Ghysels, Department of Economics, CB# 3305, Gardner Hall, Chapel Hill, NC 27599-3305.

[†]Financial Markets Department, Bank of Canada. Email: fchabiyo@bank-banque-canada.ca

[‡]Department of Economics, University of North Carolina at Chapel Hill and Department of Finance, Kenan-Flagler Business School. Email: eghysels@unc.edu

[§]Department of Economics, University of North Carolina at Chapel Hill and CIRANO, CIREQ, Montreal. Email: renault@email.unc.edu

Introduction

Several authors have recognized that pricing kernels should reward assets for their contribution to the higher order moments, such as skewness and kurtosis, of the return on the market portfolio in order to successfully explain individual asset risk premiums. The logic for including higher moments is either based on polynomial approximations to the unknown pricing kernel requiring expansions at least up to the third or fourth order (see e.g. Bansal, Hsieh, and Viswanathan (1993), Chapman (1997), among others), or else based on an expansion of the CAPM beyond the variance-return trade-off to risk sensitivities pertaining to skewness, and thus provide a rationale for the presence of higher empirical moments (see e.g. Harvey and Siddique (2000), Dittmar (2002), Jondeau and Rockinger (2006), among others).

The main drawback of pricing kernels going beyond mean-variance analysis is that they do not map directly into a fundamental preference-based asset pricing model through aggregation of individual portfolio choices. It is the purpose of this paper to close this gap. The present paper is in fact slightly more ambitious as we go beyond the usual representative agent utility models by allowing not only for heterogeneity of preferences but also for heterogeneity of beliefs among agents. In particular, we provide a general framework that maps heterogeneity of beliefs and preferences into polynomial expansions of the pricing kernel. Hence, a first contribution of our paper is to provide a structural interpretation to the empirical pricing kernels involving skewness and kurtosis. The parameter estimates obtained by Bansal, Hsieh, and Viswanathan (1993), Chapman (1997), Dittmar (2002) and Harvey and Siddique (2000), among others, can be viewed as reduced form estimates to which we attach an economic interpretation in terms of aggregation of heterogenous preferences and beliefs.

The first contribution of the paper is that we estimate pricing kernels that are functions of deep structural parameters and we show that the resulting pricing kernels fit returns very well, when appraised via traditional pricing error statistics. To handle the various sources of heterogeneity, beliefs and preferences that is, we follow the framework of Samuelson (1970) and its recent generalization of Chabi-Yo, Leisen, and Renault (2006). This generic approach allows us to derive, for risks that are infinitely small, optimal shares of wealth invested in each security that coincide with those of a Mean-Variance-Skewness-Kurtosis optimizing agent. Through these local approximations we are able to tease out the various sources of risk.

A second contribution is that our theoretical analysis reveals that the empirical pricing kernels studied by the aforementioned authors omit a component pertaining to the dispersion of investors' preferences for skewness. More specifically, the empirical pricing kernels involve the squared and the cubic market return, the right aggregates to price respectively skewness and kurtosis. However, we show that when expansions are taken to the next order required to price kurtosis, the dispersion of investors' preferences for skewness gives a nonzero weight to an additional factor ignored by studies based on a representative agent asset pricing model. The fact that heterogeneity of preferences gives rise to additional pricing factors may be related to the general theory of pricing with heterogeneity (see in particular Constantinides and Duffie (1996) and Heaton and Lucas (1995)). While the former focuses on incomplete consumption insurance, we focus instead on incompleteness with respect to nonlinear risks. Preferences for skewness leads investors to track the squared market return and the part that cannot be hedged with primitive (linear) assets introduces an additional factor of risk with a non-zero market price. The market price of this factor is proportional to the cross-sectional variance (among investors) of skewness preferences. This result is comparable to Constantinides and Duffie (1996), who find that the cross-sectional variance of consumption growth across agents gives rise to an additional priced risk factor. The advantage of our approach is that it is more appealing when it comes to empirical analysis. We can treat the cross-sectional variance of skewness preferences as a structural parameter that can be estimated from aggregate price data, whereas the cross-sectional variation of consumption growth requires consumer panel data. Finally, regarding market incompleteness it also worth noting that the small noise expansion approach of Samuelson was initially conceived to transplant insights from continuous time asset pricing models based on Brownian motions to a discrete time static investor optimization setting. In this regard, we can think of the non-hedgeable part of the squared market return as incompleteness introduced in a diffusion setting through stochastic volatility.

A third contribution of our paper is that the structural interpretation of the pricing kernel we obtain allows us to disentangle the effects of heterogeneous beliefs and preferences on asset prices. Various additional pricing factors appear in the pricing kernel with weights depending on the dispersion across investors of preferences, beliefs and interactions between them. Differences across investor's beliefs will be revealed through their different assessments of expected returns, similar to the recent literature on model uncertainty.[1] As in the model uncertainty literature, a reference probability model is used to characterize the pricing kernel. Obviously, individual investor's beliefs will differ from the reference model, and so might also be the aggregation of beliefs across agents.

A fourth contribution of the paper is that our closed form formulas for pricing kernels allow us to estimate means, variances and covariances of the distribution among investors of preference and beliefs parameters. We use this information to gauge the magnitude of aggregated belief distortions about risk premiums necessary for the pricing kernel to move inside the Hansen-Jagannathan bounds.

To sum up, we establish both theoretically and empirically a well-founded aggregation theory which tells us what are the relevant aggregates when investors are heterogeneous with respect to both preferences and beliefs.

Asset pricing models with heterogeneous beliefs and heterogeneous preferences are more realistic than representative agent models. When it comes to empirical implementation, however, we still have a long distance to go. One obvious reason is lack of data measuring heterogeneity, since data on portfolios, preferences and beliefs of individual agents is

[1]See e.g. Hansen and Sargent (2001), Anderson, Ghysels, and Juergens (2005a), Anderson, Ghysels, and Juergens (2005b), Hansen, Sargent, and Tallarini (1999), Anderson, Hansen, and Sargent (2003), Chen and Epstein (2002), Hansen, Sargent, Turmuhambetova, and Williams (2004), Kogan and Wang (2002), Liu, Pan, and Wang (2005) Uppal and Wang (2003), Maenhout (2004), among others.

very sparse and of questionable quality. The most significant progress, as far as empirical implementation is concerned, has been made with regards to empirical asset pricing with heterogeneous beliefs. Recent examples include, Shefrin (2001), Anderson, Ghysels, and Juergens (2005a) and Qu, Starks, and Yan (2003), among others. We use data similar to Anderson, Ghysels, and Juergens (2005a) which consists of returns on portfolios with strong disagreement among analysts to capture the pricing effect of heterogeneity of beliefs.

Our empirical findings, based on Hansen-Jagannathan bounds and GMM estimation of structural parameters, suggest that the heterogeneity of beliefs plays an important role in the pricing kernel and solves both the equity premium puzzle of Mehra and Prescott (1985) and the risk-free rate puzzle of Weil (1989). In the mean-variance-skewness model, we find that the skewness parameter estimates have the right sign and magnitude, yet none of the estimates are statistically significant. In contrast, the heterogeneity of beliefs parameter is strongly statistically significant with a very reasonable range of values, combined with risk tolerance parameters that are also within economically plausible ranges. When we look to the averages across industries, we find that if the return on the market in excess of the risk-free rate increases by 1%, the excess average industry return increases by roughly 0.5%. We also find that the price of risk of the beliefs portfolio is significant and positive and if the excess return on the beliefs portfolio increases by 1% the excess average industry return increases by as high as .57%, i.e. as high as the market portfolio excess return impact. The price of skewness risk of the squared market return is significant, negative and economically considerable. If the excess return on the squared market return increases by 1% the excess average industry return decreases by roughly 1.6%. It should be noted that this reduced form impact of skewness, involves not only the (statistically insignificant) skewness preference parameter, but also other structural parameters such as (variance) risk aversion. Other portfolios also appear in the reduced form pricing kernel with typically smaller, yet empirically significant impacts on returns.

The paper is organized as follows. In section 1 we start with a brief description of Samuelson's small noise expansion framework and its recent generalization of Chabi-Yo, Leisen, and Renault (2006). The key innovation of the present paper with respect to Chabi-Yo, Leisen, and Renault (2006) is the introduction of heterogeneous beliefs about expected returns. In the next section 2 we introduce a Mean-Variance-Skewness asset pricing model and derive equilibrium portfolio allocations and the implied pricing kernel with heterogeneous beliefs and preferences. Section 3 takes the analysis a step further as we study a Mean-Variance-Skewness-Kurtosis environment. In section 4 we turn our attention to the empirical analysis of our models.

1 General framework

We follow the framework of Samuelson (1970), who argued that, for risks that are infinitely small - sometimes also called small noise expansions - optimal shares of wealth invested in each security coincide with those of a mean-variance optimizing agent. Chabi-Yo, Leisen, and Renault (2006) derived a more general approximation theorem to further characterize the local sensitivity of the optimal shares with respect to other risks. For example, the Taylor expansion of the utility function one step beyond the quadratic approximation yields a price for skewness in a Mean-Variance-Skewness framework. Furthermore, one additional expansion yields a mean-variance-skewness-kurtosis approach.

The purpose of this section is to revisit the general framework of Samuelson, and expand its realm of applications. In subsection 1.1 we start from the generalization of Samuelson's result as derived by Chabi-Yo, Leisen, and Renault (2006).[2] and introduce heterogeneous preferences into Samuelson's small noise expansion setting. A final subsection 1.2 concludes with heterogeneity of beliefs. While novel in its context, our approach is very much inspired by recent work on model uncertainty.

1.1 Samuelson's Small Noise Expansions Revisited

We consider an investor s with von Neumann-Morgenstern preferences, i.e. (s)he derives utility from date 1 wealth according to the expectation over some increasing and concave function u_s evaluated over date 1 wealth. For the moment we will focus on a single investor s, be it representative or not, and later we will populate the economy with $s = 1, \ldots, S$ potentially different investors. For given risk-level σ (s)he then seeks to determine portfolio holdings $(\omega_{is})_{1\leq i\leq n} \in \mathbb{R}^n$ that maximize her/his expected utility for a given initial wealth invested q_s:

$$\max_{(\omega_{is})_{1\leq i\leq n}\in\mathbb{R}^n} Eu_s(W_s) \tag{1}$$

$$\text{with } W_s = q_s\left\{\mathbf{R}_f + \sum_{i=1}^{n}\omega_{is}\cdot(\mathbf{R}_i^s - \mathbf{R}_f)\right\}.$$

where $\mathbf{R}_f$ the gross return on the riskless asset and the solution is denoted by $(\omega_{is}(\sigma))_{1\leq i\leq n}$ and depends on the given scale of risk σ.

To define *"risk"* we turn next to the data generating process for returns. In particular, let us denote by $\mathbf{R}_i$, the (gross) return from investing one dollar in risky security $i = 1, ..., n$. The more general notation used in equation (1), namely $\mathbf{R}_i^s$, is designed to accommodate heterogeneity of beliefs which will be introduced in the next subsection. In the current subsection, where individual beliefs are not explicitly introduced, we will simply let $\mathbf{R}_i^s = \mathbf{R}_i$ for all i and s.

The random vector $\mathbf{R} = (\mathbf{R}_i)_{1\leq i\leq n}$ defines the objective joint probability distribution of interest, which is specified by the following decomposition:

$$\mathbf{R}_i(\sigma) = \mathbf{R}_f + \sigma^2 a_i(\sigma) + \sigma Y_i. \tag{2}$$

Here, $a_i(\sigma)$, $i = 1, ..., n$, are positive functions of σ. The parameter σ characterizes the scale of risk and is crucial for our analysis. In this paper we are interested in small noise expansions,

[2] See also Anderson, Hansen, and Sargent (2006) for recent work on small noise expansions.

i. e. approximations in the neighborhood of $\sigma = 0$. They provide a convenient framework to analyze portfolio holdings and resulting equilibrium allocations for a given random vector $Y = (Y_i)_{1 \leq i \leq n}$ with

$$E[Y] = 0, \text{ and } Var(Y) = \Sigma,$$

where Σ is a given symmetric and positive definite matrix. For future reference we denote by

$$\Gamma_k = E[YY^\intercal Y_k]$$

the matrix of covariances between Y_k and cross-products $Y_i Y_j$, $i, j = 1, ..., n$. Typically, asymmetry in the joint distribution of returns means that at least some matrices Γ_k, $k = 1, ..., n$ are not zero.

In equation (2), the term $\sigma^2 a_i(\sigma)$ has the interpretation of a risk premium. Obviously, equation (2) does not restrict the probability model of asset returns unless something is said explicitly about the functions $a_i(\sigma)$. Samuelson (1970) restricts $a_i(\sigma)$ to constants a_i. Under this assumption risk premiums are proportional to the squared scale of risk. Samuelson (1970) also provides a heuristic explanation for equation (2), as it closely relates to continuous time finance models. In particular, thinking in terms of Brownian motion, σ may be thought of as the square root of time, while the drift and the diffusion terms are given by the vector with components $(\mathbf{R}_f + a_i \sigma^2)$ and (σY_i) respectively.

We will show that local variations of $a_i(\sigma)$ in the neighborhood of $\sigma = 0$ allow us to characterize the price of skewness and kurtosis in equilibrium. In particular, in a second order expansion $a_i(\sigma) = a_i(0) + \sigma a_i'(0) + \sigma^2/2 a_i''(0)$, we will show that - in the case of homogeneous beliefs - specifying the price of skewness is tantamount to fixing the slope $a_i'(0)$ while the price of kurtosis is encapsulated in the curvature $a_i''(0)$. To show these results, we first need to slightly generalize the main Samuelson (1970) result about small noise expansions.

Let us reconsider investor s with von Neumann-Morgenstern preferences expressed by u_s in equation (1). Recall that the solution to the optimal asset allocation is denoted by $(\omega_{is}(\sigma))_{1 \leq i \leq n}$. The focus of interest here is the local behavior of the shares $\omega_{is}(\sigma)$ for small levels of risk, as characterized by the quantities:

$$\omega_{is}(0) = \lim_{\sigma \rightsquigarrow 0^+} \omega_{is}(\sigma) \ , \ \omega_{is}^{'}(0) = \lim_{\sigma \rightsquigarrow 0^+} \omega_{is}^{'}(\sigma) \ , \ \omega_{is}^{''}(0) = \lim_{\sigma \rightsquigarrow 0^+} \omega_{is}^{''}(\sigma) \tag{3}$$

By a slight extension of Samuelson (1970), the following holds:

A. optimal shares of wealth invested $\omega_{is}(0)$, $i = 1, ..., n$ depend on the utility function u_s only through its first two derivatives $u_s'(q_s \mathbf{R}_f)$ and $u_s''(q_s \mathbf{R}_f)$,

B. the first derivatives of optimal shares with respect to σ, $\omega_{is}'(0)$, $i = 1, ..., n$ depend on the utility function u_s only through its first three derivatives $u_s^{'}(q_s \mathbf{R}_f)$, $u_s^{''}(q_s \mathbf{R}_f)$ and $u_s^{'''}(q_s \mathbf{R}_f)$

C. the second derivatives of optimal shares with respect to σ, $\omega_{is}''(0)$, $i = 1, ..., n$ involve the first four derivatives $u_s^{'}(q_s \mathbf{R}_f)$, $u_s^{''}(q_s \mathbf{R}_f)$, $u_s^{'''}(q_s \mathbf{R}_f)$ and $u_s^{''''}(q_s \mathbf{R}_f)$.

Property (A) implies that a mean-variance approach, using a second order Taylor expansion of u_s, provides a correct characterization of optimal shares $\omega_{is}(0)$ in the neighborhood of zero risk. Moreover, property (B) implies that a third order Taylor expansion of u_s - that is a mean-variance-skewness approach - also provides a correct characterization of the slopes $\omega'_{is}(0)$ of optimal shares in the neighborhood of zero risk. Finally, property (C) implies that the curvatures $\omega''_{is}(0)$, $i = 1, ..., n$ of optimal shares involve the first four derivatives $u_s^{'}(q_s\mathbf{R}_f)$, $u_s^{''}(q_s\mathbf{R}_f), u_s^{'''}(q_s\mathbf{R}_f)$ and $u_s^{''''}(q_s\mathbf{R}_f)$. The relevant formulas are then derived from a fourth order Taylor expansion of u_s, that is a mean-variance-skewness-kurtosis approach.

Since small noise expansions are based on the local behavior of the utility function around zero risk, it is not surprising that preferences will be characterized by their derivatives of the terminal wealth in an hypothetical risk-free environment evaluated at the value $(q_s\mathbf{R}_f)$

In particular, preferences will be characterized by three parameters:

$$\tau_s = -\frac{u_s^{'}(q_s\mathbf{R}_f)}{u_s^{''}(q_s\mathbf{R}_f)} \tag{4}$$

$$\rho_s = \frac{\tau_s^2}{2}\frac{u_s^{'''}(q_s\mathbf{R}_f)}{u_s^{'}(q_s\mathbf{R}_f)} \tag{5}$$

$$\kappa_s = -\frac{\tau_s^3}{3}\frac{u_s^{''''}(q_s\mathbf{R}_f)}{u_s^{'}(q_s\mathbf{R}_f)} \tag{6}$$

called respectively risk tolerance, skew tolerance and kurtosis tolerance. The risk tolerance coefficient τ_s is positive, (note that $1/\tau_s$ is the Arrow-Pratt absolute measure of risk aversion), while the skew and kurtosis tolerance coefficients ρ_s and κ_s are assumed to be non-negative, following the literature on preferences for higher order moments (Chapman (1997), Dittmar (2002), Harvey and Siddique (2000), Jondeau and Rockinger (2006)). Note that the positivity of skew tolerance is also supported by the literature on prudence (Kimball (1990)).

Next, we show that the standard mean-variance formulas always provide the local first approximation of the demand for risky assets, irrespective of preferences for higher order moments. In particular:

Theorem 1.1 *Consider the portfolio optimization problem appearing in equation (1) and let returns be generated by equation (2). Moreover, let the preferences of investor s be specified as in (4). Then, in the limit case $\sigma \rightsquigarrow 0$, the vector $\omega_s(0) = (\omega_{is}(0))_{1\leq i\leq n}$ of shares of wealth invested by investor s is defined by:*

$$q_s\omega_s(0) = \tau_s\Sigma^{-1}a(0)$$

where $a(0) = (a_i(0))_{1\leq i\leq n}$ is the vector of risk premiums in the neighborhood of small risks.

Proof: See Appendix A.

The above result tells us that a two mutual funds theorem is valid. Besides the risk free asset, investor s chooses to hold a share of the same risky portfolio, dubbed the mean-variance mutual fund and defined by the coefficients of the vector:

$$\varsigma = \Sigma^{-1}a(0) \tag{7}$$

The individual risk tolerance coefficient τ_s defines the respective weights of the two mutual funds in investor's s portfolio. When all investors are equal, or one thinks of a single representative agent, this result coincides exactly with the standard result in case of homogeneous beliefs and joint normality of returns. In this latter case, it will be shown in the next section that the risk premium terms $a_i(\sigma)$ are in equilibrium constant functions of σ, and therefore always equal to $a_i(0)$.

In the general case of possibly skewed and leptokurtic distributed returns, the prices for skewness and kurtosis risk will play a role through the vectors of co-skewnesses and co-kurtosis of the various assets with respect to the benchmark mean-variance portfolio $\varsigma = \Sigma^{-1}a(0)$. Let us therefore define:

$$\begin{aligned} c_i &= \varsigma^\intercal \Gamma_i \varsigma = Cov[(\varsigma^\intercal Y)^2, Y_i] \\ d_i &= Cov[(\varsigma^\intercal Y)^3, Y_i] \end{aligned} \tag{8}$$

Following Kraus and Litzenberger (1976), Ingersoll (1987), Fang and Lai (1997), Harvey and Siddique (2000) and Dittmar (2002) among others, c_i (resp. d_i) is called co-skewness (resp. co-kurtosis). It is well known that the linear combination of portfolio betas yields the variance of the portfolio return. Similarly, we show that the linear combination of portfolio co-skewness and co-kurtosis yield convenient decompositions of the skewness and kurtosis of the portfolio return.

$$\begin{aligned} \sum_{i=1}^{n} \varsigma_i c_i &= E[(\varsigma^\intercal Y)^3] = E[(M - EM)^3] \\ \sum_{i=1}^{n} \varsigma_i d_i &= E[(\varsigma^\intercal Y)^4] = E[(M - EM)^4] \end{aligned}$$

where $M = \sum_{i=1}^{n} \varsigma_i \mathbf{R}_i(\sigma)$ is the return on the mean-variance mutual fund.

To conclude, it is worth showing that the role of the above coefficients c_i and d_i vanish when returns are (multivariate) Gaussian, namely: (1) an assumption of joint symmetry of the probability distribution of returns, which is in particular fulfilled in case of joint normality of returns, implies that all the matrices Γ_i and therefore all the co-skewnesses c_i are zero for all assets i, (2) joint normality of returns implies that:

$$\begin{aligned} Cov[(\varsigma^\intercal Y)^3, Y_i] &= Cov[(\varsigma^\intercal Y)^3, E(Y_i|\varsigma^\intercal Y)] \\ &= E[(\varsigma^\intercal Y)^4] Var^{-1}[\varsigma^\intercal Y] Cov[Y_i, \varsigma^\intercal Y] \end{aligned}$$

and therefore:

$$d_i = 3Var[\varsigma^\intercal Y] Cov[Y_i, \varsigma^\intercal Y] = 3Var[\varsigma^\intercal Y] a_i(0) \tag{9}$$

Hence, the reason why the co-kurtosis coefficients do not play any role in case of joint normality is simply because they are proportional to usual the beta coefficients or, equivalently, to the CAPM risk premium terms.

1.2 Heterogeneous beliefs

In this section we discuss heterogenous beliefs and do so by starting with characterizing beliefs-distorted risk premiums, followed by beliefs-mimicking portfolios.

1.2.1 Beliefs-distorted risk premiums

We noted that equation (2) closely relates to continuous time finance models. A key insight we add to Samuelson's original setting is inspired by the recent work on model uncertainty.[3] This literature has adopted the view that alternative models, not far from an assumed reference model in terms of entropy, are absolutely continuous with respect to the reference model. This implies, by the Girsanov theorem, that (local) alternatives differ only in terms of drift functions.[4]

The analogy with the model uncertainty literature leads us to consider an investor s as having personal beliefs differing from the reference model (2) by a s-specific factor distorting the risk premium $\sigma^2 a_i(\sigma)$. Obviously, this factor should not intervene in the limit case of homogeneous beliefs ($\sigma = 0$). Hence, the beliefs of investor s beliefs, for $s = 1, ...S$, are defined by the following stochastic model for returns:

$$\mathbf{R}_i^s(\sigma) = \mathbf{R}_f + \sigma^2(a_i(\sigma) + \sigma h_{is}(\sigma)) + \sigma Y_i \tag{10}$$

where the function $\sigma h_{is}(\sigma)$ represents beliefs distortions with respect to the *"objective"* risk premium $a_i(\sigma)$. Note that the notion of objective risk premiums will only be well defined once we implement the model empirically, when the econometrically identified pricing kernel will be defined with respect to the so called *"objective"* probability measure. The latter probability measure features returns as in equation (2) where all beliefs distortions have vanished. It should be noted however, that we do not maintain the assumption that the *"average"* investor's beliefs correspond to the objective probability. This is an issue, notably discussed at length in Anderson, Ghysels, and Juergens (2005a). As the latter point out, the assumption that agents are on average correct is one of weak rational expectations, and often rejected in the behavioral finance literature (see Anderson, Ghysels, and Juergens (2005a) for further discussion).

We do impose some constraints, however, with respect to the distortions of beliefs of the average investor. In particular, note that in equation (10) the beliefs distortion function $h_{is}(\sigma)$ is pre-multiplied by σ, and this ensures that the price of risk $a_i(0)$ is not modified by beliefs distortions in the neighborhood of zero risk. Consequently, heterogeneity of beliefs will not be identified through asset demands and equilibrium prices in the limit case as σ

[3] See for instance, Hansen and Sargent (2001), Anderson, Ghysels, and Juergens (2005a), Anderson, Ghysels, and Juergens (2005b), Hansen, Sargent, and Tallarini (1999), Anderson, Hansen, and Sargent (2003), Chen and Epstein (2002), Hansen, Sargent, Turmuhambetova, and Williams (2004), Kogan and Wang (2002), Liu, Pan, and Wang (2005) Uppal and Wang (2003), Maenhout (2004), among others.

[4] Note also that according to Maenhout (2004), this restriction is entirely natural for the portfolio problem we are interested in, as a preference for robustness is often motivated by substantial uncertainty about the expected return, and therefore precisely the drift term.

goes to zero. Therefore, in our framework, heterogeneity of beliefs is unrevealed within the context of the standard mean-variance analysis. In fact, the mean-variance mutual fund result (1.1) remains valid within the more general setting with heterogenous beliefs:

Theorem 1.2 *Consider the portfolio optimization problem appearing in equation (1) and let returns be generated by equation (10). Moreover, let preferences be specified as in (4). Then, in the limit case $\sigma \rightsquigarrow 0$, the vector $\omega_s(0) = (\omega_{is}(0))_{1\leq i\leq n}$ of shares of wealth invested is defined by:*

$$q_s\omega_s(0) = \tau_s\Sigma^{-1}a(0)$$

where $a(0) = (a_i(0))_{1\leq i\leq n}$ is the vector of risk premiums in the neighborhood of small risks.

Proof: See Appendix A.

1.2.2 Beliefs-mimicking portfolios

Theorem 1.2 implies that the portfolio and pricing effects of heterogeneous beliefs will only appear through higher order moment expansions of the asset pricing kernel, when preferences for respectively high positive skewness and possibly low kurtosis are taken into account.

It is worth discussing at this stage the main implications of the modelling strategy we pursue as it transpires in Theorem 1.2. A first issue pertains to identification. In particular, distortions of beliefs could be confounded with preferences for skewness in the slopes of optimal portfolio shares, i.e. $\omega'_{is}(0)$. Hence, one may wonder whether both effects can be disentangled in the return equation (10) since both give rise to the price of risk being affected by a function of σ. We will show in section 2 that we can clearly separate the two effects - beliefs distortions and preferences for skewness - in the risk premium per unit of variance. More specifically, preference for skewness will appear in the vector $c = (c_i)_{1\leq i\leq n}$ of co-skewnesses scaled by skew tolerance, whereas distortion of beliefs, non-uniform among the various assets, will give rise to additional mutual funds, we will call *"beliefs portfolios "*. To see this, note from equation (10) that the slope of the risk premium in the neighborhood of zero risk is decomposed as $(a'_i(0) + h_{is}(0))$, where $a'_i(0)$ is the slope (at zero risk) of the objective (i.e. independent of s) risk premium function.

In the initial setup of Samuelson (1970) the slope $a'_i(0)$ was set to zero. The role of the slope $a'_i(0)$, was first exploited in Chabi-Yo, Leisen, and Renault (2006), to price the higher-order factors of risk, the price of skewness in their case, not captured by the mean-variance portfolio. In this paper we extend this approach and further decompose the slope into two components. The first component will price skewness, as originally suggested in Chabi-Yo, Leisen, and Renault (2006), whereas the second component will related to beliefs portfolios, with the purpose to hedge the heterogeneity of beliefs represented via the distortions $h_{is}(0)$. In particular, we can define:

Definition 1.1 *A beliefs-distortion function $h_{is}(\sigma)$ according to (10) gives rise to the following investor's beliefs portfolios:*

$$\varsigma_b^{(s)} = \Sigma^{-1}[\alpha_{\bullet s} \odot a(0)] \tag{11}$$

where $\odot$ is the Hadamard (element-wise) product of matrices and $\alpha_{\bullet s} \equiv (\alpha_{is})_{(1 \leq i \leq n)}$, where α_{is} is defined such that: $h_{is}(0) = \alpha_{is} a_i(0)$.

Note that the above definition is very general and only involves a mild implicit assumption. Namely, beliefs distortions are only about risk premiums that the traditional CAPM does not set to zero. In particular, since the vector $a(0)$ of prices of risk will be conformable to a standard CAPM beta-pricing, we expect that a zero-beta asset i will be endowed not only with a zero objective risk premium $a_i(0)$ but also a zero beliefs distortion $h_{is}(0)$ for any investor s.

In the next section it will be shown that beliefs portfolios matter when there is at least some investor s, with $(h_{is}(0))_{(1 \leq i \leq n)}$ not proportional to $(a_i(0))_{(1 \leq i \leq n)}$, where the latter determines the mean-variance portfolio. As a matter of fact, a beliefs portfolio $\varsigma_b^{(s)}$ does not coincide (up to a scaling factor) with the mean-variance portfolio $\varsigma = \Sigma^{-1} a(0)$, if and only if the components of $\alpha_{\bullet s}$ are not all equal. Hence, whenever for some agent s, the (α_{is}) are not equal, will imply that we need beyond the mean-variance portfolio also *"beliefs representing portfolios "* where the assets are re-weighted in proportion to the associated beliefs distortions, or *"fad effects"*.

The above discussion prompts the question whether we need S linearly independent portfolios, as many as there are agents. If $(J+1)$ is the dimension of the subspace $\mathbf{R}^n$ spanned jointly by $(\alpha_{\bullet s})_{1 \leq s \leq S}$ and the n-dimensional sum vector $\mathbf{1} = (1, \ldots, 1)'$, we will have a J-dimensional structure of beliefs distortions. This will lead to J mutually linearly independent vectors $\alpha^j \equiv (\alpha_i^j)_{1 \leq i \leq n}$, also independent of $\mathbf{1}$. It will be shown in the next section that this beliefs heterogeneity will yield J beliefs-mimicking mutual funds:

$$\varsigma_b^j = \Sigma^{-1}[\alpha^j \odot a(0)] \qquad j = 1, \ldots, J \tag{12}$$

Without loss of generality we will henceforth normalize all the weights $\Sigma^{-1}[\alpha^j \odot a(0)]$ such that:

$$\frac{1}{n} \sum_{i=1}^{n} \alpha_i^j = 1 \quad \forall j = 1, \ldots, J$$

This normalization ensures comparability with mean-variance analysis, as $\varsigma = \Sigma^{-1}[\mathbf{1} \odot a(0)]$ $= \Sigma^{-1} a(0)$. We can also unambiguously define a system of belief loadings $(\lambda_{sj})_{(1 \leq j \leq J)}$ for each investor s :

$$\begin{aligned} \alpha_{\bullet s} &= \lambda_{s0} \mathbf{1} + \sum_{j=1}^{J} \lambda_{sj} \alpha^j \qquad (13) \\ &= (\lambda_{s0} + \sum_{j=1}^{J} \lambda_{sj}) \mathbf{1} + \sum_{j=1}^{J} \lambda_{sj} (\alpha^j - \mathbf{1}) \end{aligned}$$

Large loadings λ_{sj} for some j imply that investor s strongly disagrees with $a(\sigma)$ being the correct description of the cross-section of expected returns. (S)he rather believes that some asset i will have relatively higher (resp. lower) expected return when α_i^j is larger (resp. smaller) than one. It is worth noting that when all the non-zero components of α^j are equal, an interesting case emerges as $\alpha^j \odot a(0)$ amounts to a situation where some risk premiums are zero while all other components are multiplied by the same re-normalization factor. Moreover, a diagonal matrix Σ yields a beliefs portfolio $\Sigma^{-1}[\alpha^j \odot a(0)]$ proportional to the mean-variance portfolio, except that the portfolio shares for some particular assets are annihilated and this because there is no need to introduce a discrepancy between beliefs and objective value of the risk premiums. This approach will be pursued in the empirical section 4 while neglecting possible interactions between correlations and beliefs.

It is worth reminding the reader of the fact that the above analysis pertains to *slopes in the neighborhood of zero risk,* e.g. the slopes of optimal portfolio shares $\omega_s'(0) = (\omega_{is}'(0))_{1 \leq i \leq n}$. We could however also consider higher order derivatives, like the curvature $\omega_s''(0)$. As far as first order derivatives are concerned, they will capture discrepancies, due to both preferences for skewness and heterogeneity of beliefs, with respect to the common mean-variance model. When considering higher order derivatives, like $\omega_s''(0)$, we will be able to capture preferences for kurtosis as well as higher order effects associated with heterogeneity of beliefs about expected returns. One may therefore expect additional mutual funds to emerge, designed to hedge not only against kurtosis, but also against heterogeneity of beliefs about co-skewness coefficients, or co-skewness not only with respect to the mean-variance portfolio, but also beliefs-representing portfolios.

To avoid a proliferation of mutual funds, we will make some simplifying assumption in the mean-variance-skewness-kurtosis asset pricing model introduced in this paper. The key simplifying assumption will be that:

$$h_{is}'(0) = \alpha_{is} a_i'(0). \tag{14}$$

Without this simplification, we would end up with different types of beliefs portfolios, such as $\Sigma^{-1}[\alpha_{\bullet s} \odot a(0)]$ and $\Sigma^{-1}[\tilde{\alpha}_{\bullet s} \odot a'(0)]$, the latter because $h_{is}'(0) = \tilde{\alpha}_{is} a_i'(0)$. Restricting $\tilde{\alpha} = \alpha$, as in equation (14) is justifiable as simplification. A sufficient condition is the homogeneity condition:

$$h_{is}(\sigma) = \alpha_{is} a_i(\sigma), \tag{15}$$

which corresponds to the following stochastic model for asset returns:

$$R_i^s(\sigma) = R_f + \sigma^2 a_i(\sigma)[1 + \sigma \alpha_{is}] + \sigma Y_i \tag{16}$$

While the above data generating process is restrictive with regards to equation (10), it is worth emphasizing this only applies to the curvature $\omega_s''(0)$. As far as the mean-variance-skewness-beliefs asset pricing model is concerned, we do *not* require the homogeneity restriction appearing in (15). For instance, with a one-dimensional beliefs distortion (i.e. J =1):

$$\alpha_{\bullet s} = \lambda_{s0} \mathbf{1} + \lambda_s \alpha$$

several additional effects of beliefs distortions are expected in the curvatures $\omega_{is}''(0)$ of optimal portfolio shares. Skipping for the moment the details discussed later in the paper, we will show that beyond the vector of co-kurtosis, higher order small noise expansions will result in three additional portfolios involving beliefs distortions:

- A portfolio $\varsigma_{bb} = \Sigma^{-1}[\alpha \odot \alpha \odot a(0)]$ which represents something like a beliefs portfolio built on the basic beliefs portfolio $\varsigma_b = \Sigma^{-1}[\alpha \odot a(0)]$ defined above. Obviously, with uniform beliefs distortion ($\alpha_i = 1$ for all i), all portfolios coincide with the basic mean-variance mutual fund $\varsigma = \Sigma^{-1}a(0)$.

- A portfolio based on a cross-co-skewness measure between the mean-variance mutual fund ς and the beliefs portfolio ς_b :

$$c_{ib}= \varsigma^{\intercal}\Gamma_i\varsigma_b= Cov[(\varsigma^{\intercal}Y)(\varsigma_b^{\intercal}Y), Y_i] \tag{17}$$

  The cross-co-skewnesses c_{ib} are all zero like the co-skewnesses c_i when the primitive asset return distributions are symmetric ($\Gamma_i = 0$). For general distributions, they coincide in the case of uniform beliefs distortion ($\alpha_i = 1$ for all i).

- A portfolio based on beliefs about co-skewnesses, that is based on the vector $\alpha \odot c$.

2 Mean-Variance-Skewness-Beliefs Pricing

In this first of two sections we go beyond the standard mean-variance formulas that form the basis for the CAPM. In a first subsection 2.1 we analyze the individual investor's problem, before deriving in the next subsection 2.2 the implications for equilibrium allocations and prices. A final subsection 2.3 derives the pricing kernel.

2.1 The individual investor problem

A Taylor expansion of the utility function one order beyond the quadratic approximation characterizes the demand for additional portfolios beyond the basic mean-variance mutual fund $\varsigma = \Sigma^{-1}a(0)$. Namely, two additional portfolios appear, as stated in the following theorem:

Theorem 2.1 *Assume the setting of Theorem 1.2, with preferences specified as in (4) and (5), where ρ_s is the skewness tolerance. Then, in the neighborhood of $\sigma = 0$, the first order approximation $[\omega_s(0) + \sigma\omega_s'(0)]$ of the vector $\omega_s(\sigma)$ of shares of wealth invested is defined by:*

$$q_s[\omega_s(0) + \sigma\omega_s'(0)] = \tau_s\Sigma^{-1}[(\mathbf{1} + \sigma\alpha_{\bullet s})] \odot a(0) + \sigma\tau_s\Sigma^{-1}[\rho_s c + a'(0)]. \tag{18}$$

Proof: See Appendix B.

Recall that identification issues may potentially arise, as one may expect that distortions of beliefs may be confounded with preferences for skewness in the slopes of optimal portfolio shares, i.e. $\omega'_{is}(0)$. Theorem 2.1 clearly disentangles the respective roles of preferences for skewness and distortion of beliefs. Besides the mean-variance portfolio $\varsigma = \Sigma^{-1}a(0)$, two additional portfolios are included in the demand of investor s : (1) the beliefs-distorted portfolio $\varsigma_b^{(s)} = \Sigma^{-1}[\alpha_{\bullet s} \odot a(0)]$ and (2) the so-called skewness portfolio (see Chabi-Yo, Leisen, and Renault (2006)) defined by $\varsigma_{sk} = \Sigma^{-1}c$. Note that the coefficient of the beliefs-distorted portfolio in the investor's demand does not involve anything related to skewness or skewness preferences. Likewise, the coefficient of the skewness portfolio in the demand of investor s does not involve anything related to beliefs distortion and is simply proportional to the intensity ρ_s of skew tolerance.

It was previously noted that the coefficient c_i measures the contribution of asset i to the skewness of the mean-variance portfolio ς and can be called the co-skewness of asset i in the portfolio. This is the reason why, in the particular case of a diagonal covariance matrix Σ, a large co-skewness c_i will increase the demand for asset i, particularly when investor s has a strong preference for positive skewness, as measured by ρ_s. Therefore, individual preferences for positive skewness will increase, ceteris paribus and up to correlation effects, the equilibrium price of assets with positive co-skewness. This effect will appear in the equilibrium value $a'(0)$ of risk premium slopes.

An alternative interpretation of the skewness portfolio follows from Chabi-Yo, Leisen, and Renault (2006) who observe that the affine regression of the squared return $(\varsigma'\mathbf{R})^2$ of the mean-variance mutual fund on the vector $\mathbf{R} = (\mathbf{R}_i)_{1\leq i\leq n}$ of returns on risky assets is an affine function of ς_{sk}. Hence, while derivative assets with nonlinear payoffs may be in practice a way to trade skewness, the skewness portfolio ς_{sk} appears in our framework as the best way to replicate the nonlinear payoff $(\varsigma'\mathbf{R})^2$ via trading exclusively assets which have payoffs that are linear functions of primitive returns.

2.2 Equilibrium Prices and Agent Demands

We turn now to equilibrium prices and demand, starting with some assumptions about aggregate quantities:

Assumption 2.1 *We assume that the net supply of each risky asset $i = 1, ..., n$ is exogenous and independent of the scale of risk σ. Then the Taylor expansions of individual portfolios shares must fulfill the following market clearing conditions:*

$$\sum_{s=1}^{S} q_s\omega_s(0) = S\overline{\omega}, \qquad \sum_{s=1}^{S} q_s\omega_s^{'}(0) = 0$$

where S is the number of (types of) agents in the economy.

Assumption 2.2 *There is a J-dimensional structure of beliefs distortions:*

$$\alpha_{\bullet s} = \lambda_{s0}\mathbf{1} + \sum_{j=1}^{J} \lambda_{sj}\alpha^j \qquad \forall s = 1, \ldots, S. \tag{19}$$

Note that the beliefs structure in Assumption 2.2 implies that the first order expansion of the vector of risk premiums can be written as:

$$ER^s(\sigma) - R_f\mathbf{1} = \sigma^2[(1+\sigma\lambda_{s0})\mathbf{1} + \sigma a'(0)] \quad (20)$$

$$+\sigma^2[\sigma \sum_{j=1}^{J} \lambda_{sj}\alpha^j \odot a(0)], \quad (21)$$

With a slight abuse of language, we will refer to a 0-dimensional structure when $\alpha_{\bullet s} = \lambda_{s0}\mathbf{1}$, and thus the risk premium expansion is the first line of equation (20). In this case, a positive λ_{s0} implies an overconfident investor s who uniformly scales expectations with an upward bias (relative to the objective expectations) across all assets.

We will show that Assumptions 2.1 and 2.2 implies that the *"market portfolio"* $\overline{\omega}$ is the portfolio selected by the average investor, with average initial wealth $\overline{q} = 1/S\sum_{s=1}^{S} q_s$ and average preferences and beliefs. To do so, we need the following quantities:

Definition 2.1 *The average investor is characterized by the following population averages:*

$$\overline{\tau} = \frac{1}{S}\sum_{s=1}^{S}\tau_s, \quad \overline{\rho} = \frac{\sum_{s=1}^{S}\tau_s\rho_s}{\sum_{s=1}^{S}\tau_s}, \quad \overline{\lambda}_j = \frac{\sum_{s=1}^{S}\tau_s\lambda_{sj}}{\sum_{s=1}^{S}\tau_s} \qquad j = 0, 1, \ldots, J \quad (22)$$

Note that the average skew tolerance and average loadings of beliefs distortions are computed with weights proportional to risk tolerance. Hence:

$$\sum_{s=1}^{S}\tau_s(\rho_s - \overline{\rho}) = 0, \qquad \sum_{s=1}^{S}\tau_s(\lambda_{sj} - \overline{\lambda}_j) = 0 \qquad j = 0, 1, \ldots, J$$

As noted before, it is important to remind the reader that we did not assume that the average investor's beliefs coincide with the expectations under the objective probability model (2). Hence, we do not impose that the averaged loadings of beliefs distortions $\overline{\lambda}_j$ are all zero.

We substitute $\omega_s(0)$, $\omega_s'(0)$, as characterized by Theorems 1.2 and 2.1, into the market clearing condition and obtain:

$$a(0) = \Sigma\overline{\omega}/\overline{\tau} \qquad a'(0) = -\overline{\rho}c - \overline{\lambda}_0(\alpha \odot a(0)) - \textstyle\sum_{j=1}^{J}\overline{\lambda}_j\alpha^j \odot a(0) \quad (23)$$

To summarize, we have the following theoretical result:

Theorem 2.2 *Let Assumption 2.1 and 2.2 and Definition 2.1 hold and let the preferences be as in Theorem 2.1. Then the first order approximation of the asset demand of investor s in equilibrium is:*

$$q_s[\omega^s(0)+\sigma\omega_s'(0)] = \frac{\tau_s}{\overline{\tau}}\left\{[1+\sigma(\lambda_{s0}-\overline{\lambda}_0)]\overline{\omega} + \sigma\Sigma^{-1}\sum_{j=1}^{J}(\lambda_{sj}-\overline{\lambda}_j)[\alpha^j \odot \Sigma\overline{\omega}]\right\} + \tau_s(\rho_s - \overline{\rho})\Sigma^{-1}\sigma c$$

Theorem 2.2 is a mutual funds separation theorem which displays $(J+2)$ mutual funds in equilibrium:

- Similar to the standard Sharpe-Lintner CAPM, investor s holds a share the of market portfolio $\overline{\omega} = \overline{\tau}\varsigma$ proportional to the mean-variance mutual fund. Per unit of wealth invested, the size of this share is determined by the risk tolerance of investor s relative to the average one.

- A non-zero vector c of co-skewnesses appears when some asset return distributions are skewed. This vector c gives rise to an additional mutual fund defined by shares proportional to $\varsigma_{sk} = \Sigma^{-1}c$. The corresponding portfolio is held in a positive quantity by investors s whose skewness tolerance ρ_s is higher than average.

- A J-dimensional structure of beliefs distortions with heterogeneous beliefs may even introduce J additional beliefs portfolio defined by shares proportional to $\varsigma_b^j = \Sigma^{-1}[\alpha^j \odot a(0)]$. It is held in a positive quantity by investors s whose expectations on risk premiums are scaled by a beliefs loadings λ_{js} higher than average. Its composition deviates from the market portfolio $\bar{\omega}$ the more the beliefs coefficient α_i^j for asset i deviates from one.

Note that, by definition, the skewness portfolio ς_{sk} and the beliefs portfolios ς_b^j, $j = 1, \ldots, J$, are all in zero net aggregate supply.

A simple case of the additional mutual fund comes with a 0-dimensional beliefs distortion structure. In this case, heterogeneity of beliefs does not really give rise to an additional mutual fund since the beliefs portfolio coincides with the market portfolio. Up to the skewness portfolio, individual asset demands are only shares of the market portfolio according to the formula:

$$(\tau_s/\overline{\tau})[\mathbf{1} + \sigma(\lambda_{0s} - \overline{\lambda}_0)]\overline{\omega} \tag{24}$$

Hence, the only role of heterogeneity of beliefs in this case is an apparent distortion of risk aversion. In particular, consider an overconfident investor s, as characterized by a distortion factor λ_{0s}, scaling uniformly above average her expectations over all assets returns. In terms of asset demands, such an investor will be observationally equivalent to an investor with an average distortion of beliefs but a risk tolerance larger than τ_s by a factor of $[\mathbf{1}+\sigma(\lambda_{0s}-\overline{\lambda}_0)]$.

The above result coincides with similar findings in the model uncertainty or robustness literature, see for instance Maenhout (2004) (see his formula (17), page 962). Therefore, as concern for robustness amounts to an increase in effective risk aversion, we conclude that overconfidence results in a decrease of effective risk aversion.

Our setting also nests Uppal and Wang (2003) who allow for non-uniform concerns for robustness among risky assets or equivalently for asset-dependent distortions of expected returns. In a similar way they end up with effective risk aversions different for each asset (see their formula (28), page 2476). Moreover, we have shown that the various factors α_i^j scale the various components of the market return differently (see the difference between $\overline{\omega} = \overline{\tau}\varsigma = \overline{\tau}\ \Sigma^{-1}a(0)$ and $\varsigma_b^j = \Sigma^{-1}[\alpha^j \odot a(0)]$), giving rise to additional mutual funds.

It is reasonable to assume that there is no distortion for some assets, hence expected returns are agreed upon by all agents ($\alpha_i^j = 0$ for such an asset i) while expected returns for other assets are uniformly uncertain. This gives rise to a beliefs portfolio $\varsigma_b^j = \Sigma^{-1}[\alpha^j \odot a(0)]$ where only the uncertain assets are included. This feature will be exploited in the empirical analysis reported later in the paper.

To conclude, as far as preferences for robustness is concerned, it is worth noting that in a Gaussian framework, our portfolio and asset pricing model is observationally equivalent to a general version of both Uppal and Wang (2003) and Maenhout (2004). However, we need to emphasize an important difference in case of significant preferences for positive skewness and asymmetries in asset payoffs as well. While such asymmetries and skewness preferences have been well documented (see Chabi-Yo, Leisen, and Renault (2006) and references therein), we argue that they must be jointly identified with heterogeneity of beliefs or concern for robustness with regards to model uncertainty. They both correspond to higher order terms in the risk-return trade-off and, for this reason, must be considered simultaneously. However, since they give rise to different mutual funds, it leaves room for separate identification of these two effects, as shown in the next subsection and in our empirical study as well.

2.3 The Pricing Kernel

The empirical work discussed in section 4 will involve asset pricing data, as is typically the case, as there is no data on asset demands. It is therefore necessary to characterize our model in terms of implied pricing kernels. This will allows us to draw comparisons with the empirical work of Bansal, Hsieh, and Viswanathan (1993), Chapman (1997), Harvey and Siddique (2000), Dittmar (2002), among others.

The fundamental asset pricing equation tells us that the pricing kernel m should be able to price correctly all assets at hand and:

$$Em = 1/R_f \qquad \text{and} \qquad Em\mathbf{R} = \mathbf{1} \tag{25}$$

where expectations are computed with respect to the objective probability measure defined by equation (2), namely:

$$\mathbf{R} = \mathbf{R}_f\mathbf{1} + \sigma^2 a(\sigma) + \sigma Y$$

As usual, we may expect the pricing kernel to be spanned by the returns on the mutual funds which appear in individual equilibrium asset demands. Therefore, when the risk premium vector is approximated by its first order expansion $\sigma^2[a(0) + \sigma a'(0)]$, we may expect the pricing kernel to be spanned by (i) the risk free rate $\mathbf{R}_f$, (ii) the market return $\mathbf{R}_M = \overline{\omega}^\intercal\mathbf{R}$ $= \overline{\tau}\varsigma^\intercal\mathbf{R}$, (iii) the payoffs on the J beliefs portfolio $\mathbf{R}_{Mb}^j = \overline{\tau}(\varsigma_b^j)^\intercal\mathbf{R}$, $j = 1, \ldots, J$, and (iv) the payoff on the skewness portfolio $\mathbf{R}_{Msk} = \varsigma_{sk}^\intercal\mathbf{R}$. Moreover, we noted that the affine regression of the squared market return $\mathbf{R}_M^2$ on the vector $\mathbf{R}$ of primitive returns is an affine function of the payoff of the skewness portfolio. Therefore, we do not modify the fundamental pricing relationship (25) when we replace the payoff $\mathbf{R}_{Msk}$ of the skewness portfolio with $\mathbf{R}_M^2$ in the pricing kernel m. This is the reason why we can prove the following result:

Theorem 2.3 *The vector of asset risk premiums $\sigma^2 a(\sigma)$ approximated by $\sigma^2[a(0)+\sigma a'(0)]$, (with $a(0)$ and $a'(0)$ determined in equation (23)) is consistent with the pricing kernel:*

$$m(\sigma) = A_0 + A_1 md(\mathbf{R}_M) + \sum_{j=1}^{J} A_{b,j}^{j} md(\mathbf{R}_{Mb}^{j}) + A_3 md\{[md(\mathbf{R}_M)]^2\}$$

with: $A_0 = 1/\mathbf{R}_f$, $A_1 = (-1/(\mathbf{R}_f\overline{\tau}))(1-\sigma\overline{\lambda}_0)$, $A_3 = \overline{\rho}/(\mathbf{R}_f\overline{\tau}^2)$, $A_{b,j} = (\sigma\overline{\lambda}_j)/(\mathbf{R}_f\overline{\tau})$, *$j$ = 1, ..., J, $md(X) = X - E(X)$ denotes the mean-deviation part of a random variable X, and $md([md(X)]^k)$ corresponds to $(X-E(X))^k$ - $E(X-E(X))^k$. In addition, the market return $\mathbf{R}_M = \overline{\omega}^\intercal R = \overline{\tau}\varsigma^\intercal \mathbf{R}$, the payoffs on the J beliefs portfolios $\mathbf{R}_{Mb}^{j} = \overline{\tau}(\varsigma_b^j)^\intercal R$, j = 1, ..., J.*

Proof: See Appendix C.

It is worth noting that in Theorem 2.3, the small noise expansions enabled us to aggregate individual preferences and beliefs, which yielded the factors for equilibrium pricing. For example, compared to the existing literature it was not a priori assumed that the squared market return was the relevant aggregate for pricing skewness. Indeed, in the prior literature (see e.g. Harvey and Siddique (2000) and Dittmar (2002)), the squared market return appears through Taylor expansions of a representative agent utility function. In contrast, in Theorem 2.3 the squared market return appears as a pricing factor for co-skewness whose mimicking portfolio is the skewness portfolio. The latter has been derived as the correct aggregate, that is the mutual fund held in equilibrium by all investors whose skewness tolerance is higher than average. Similarly, the beliefs portfolios appear as mutual funds all heterogeneous investors will use to hedge their differences in beliefs, while they were previously put forward by Anderson, Ghysels, and Juergens (2005a) only as a reduced form approximation.

It is worth noting that a beliefs portfolio is relevant in the pricing kernel if and only if the average investor has a non-zero belief distortion. Hence, in the particular case where aggregation of beliefs would generate the objective probability distribution, beliefs distortions have no pricing effect at this order of expansion.

When all investors are considered as identical with the same initial endowment $\overline{q}$, then the aggregation issue becomes straightforward: they all chose the same portfolio $\omega(\sigma)$ (in particular $\overline{\omega} = \overline{q}\,\omega(0)$) and the relevant averages $\overline{\tau}$, $\bar{\rho}$, and $\bar{\lambda}_j$ are interpreted as characteristics of a representative investor identical to all the investors in the economy:

$$\begin{aligned} \overline{\tau} &= -\frac{u'(\overline{q}\mathbf{R}_f)}{u''(\overline{q}\mathbf{R}_f)} \\ \overline{\rho} &= \frac{\overline{\tau}^2}{2}\frac{u'''(\overline{q}\mathbf{R}_f)}{u'(\overline{q}\mathbf{R}_f)} \end{aligned}$$

In such a case: (i) the coefficient A_1 of the market return is $(\mathbf{R}_f^{-1}\overline{\tau})$ is the same as in the common CAPM formulas, (ii) the coefficient A_2 of the squared market return is $(\bar{\rho}/\mathbf{R}_f\overline{\tau}^2)$ as in Dittmar (2002), (iii) the coefficient $A_{b,j}$ of the return on the beliefs portfolio ς_b^j is $(\sigma/\mathbf{R}_f)(\bar{\lambda}_j/\overline{\tau})$ because in this case, a common degree of for instance overconfidence among investors for a subset of assets (positive $\bar{\lambda}_j$ common to all investors) is observationally equivalent to an effective risk aversion different for each asset as in Uppal and Wang (2003), Anderson, Ghysels, and Juergens (2005a), among others.

To summarize, we established a well-founded aggregation theory which tells us what are the relevant aggregates when investors are heterogeneous with respect to both preferences and beliefs. This is in contrast to the results of Gorman (1953) who used exact aggregation argument, whereas we use Samuelson's small noise expansion setting. Moreover, the average preference for skewness must be taken into account to assess the pricing effect of the average distortion of beliefs. Most importantly, at *order zero*, our expansions reduce to the common CAPM.

Finally, if we want to characterize the pricing effects of the dispersion among investors of both preferences and beliefs, that is to exhibit some risk factors whose price depends not only on average values $\bar{\rho}$ and $\bar{\lambda}$ of skew tolerance and distortion of beliefs respectively but also on their cross-sectional dispersion among investors, we need to go one step further, that is to take also into account the average tolerance for kurtosis, the topic of the next section.

3 Mean-Variance-Skewness-Kurtosis-Beliefs Pricing

Having introduced the Mean-Variance-Skewness asset pricing model, we now turn to its further extension, that is, we add kurtosis risk. The structure of the section is the same as the previous one. In a first subsection 3.1 we analyze the individual investor problem, before deriving in the next subsection 3.2 the implications for equilibrium allocations and prices. A final subsection 3.3 derives the pricing kernel.

3.1 Individual Asset Demand

A Taylor series expansion of the utility function two steps beyond the quadratic approximation yields a price for skewness and kurtosis, or a mean-variance-skewness-kurtosis framework using the simplified model (16) for returns. The first order conditions are identical to the mean-variance-skewness agent first order conditions, namely:

$$E[u_s^{'}(W^s(\sigma))(\sigma a_i(\sigma)(1 + \alpha_{is}\sigma) + Y_i)] = 0$$

This means that the level $\omega_s(0)$ and slope $\omega_s^{'}(0)$ of the portfolio weights of agent s are identical to those obtained in Theorem 2.1. To further characterize the curvature $\omega_s^{''}(0)$ of the portfolio weight of agent s we rely on the extended Samuelson (1970) result, namely we use a Taylor expansion of the utility function up to the fourth degree:

Theorem 3.1 *Assume the setting of Theorem 1.2, with preferences be specified as in (5), where ρ_s and κ_s denote respectively the skewness and kurtosis tolerances. Then, in the* *neighborhood of $\sigma = 0$, the second order approximation $[\omega_s(0) + \sigma\omega_s^{'}(0) + \sigma^2\omega_s''(0)/2]$ of the vector $\omega_s(\sigma)$ of shares of wealth invested is defined by:*

$$
\begin{aligned}
\omega_s(0) &= (\tau_s/q_s)\varsigma \\
\omega_s'(0) &= (\tau_s/q_s)[\rho_s\varsigma_{sk} + \varsigma_b^s + \Sigma^{-1}a'(0)] \\
\omega_s''(0)/2 &= (\tau_s/q_s)[\varsigma_b^{s\prime} + \Sigma^{-1}a''(0)/2] \\
&\quad -\kappa_s(\tau_s/q_s)\varsigma_{kurt}/2 + (\tau_s/q_s)(3\rho_s - 1)(\varsigma^\intercal\Sigma\varsigma)\varsigma \\
&\quad +2(\tau_s/q_s)\Sigma^{-1}\rho_s\left[\rho_s(\varsigma^\intercal\Gamma_i\varsigma_{sk}) + \varsigma^\intercal\Gamma_i\varsigma_b^s + \varsigma^\intercal\Gamma_i\varsigma'\right]_{1\le i\le n}
\end{aligned} \tag{26}
$$

where:

$$
\begin{aligned}
\varsigma &= \Sigma^{-1}a(0) & \varsigma_{sk} &= \Sigma^{-1}c \\
\varsigma_{kurt} &= \Sigma^{-1}d & \varsigma_b^s &= \Sigma^{-1}[\alpha_{\bullet s}\odot a(0)] \\
\varsigma' &= \Sigma^{-1}a'(0) & \varsigma_b^{s\prime} &= \Sigma^{-1}[\alpha_{\bullet s}\odot a'(0)]
\end{aligned} \tag{27}
$$

Proof: See Appendix D.

We noted in the previous section that the formulas for asset demand allow us to disentangle the effects of preferences versus distortion of beliefs. In this respect, Theorem 3.1 introduces again several distinct portfolios:

First, the mean-variance portfolio $\varsigma = \Sigma^{-1}a(0)$ and is replaced by a similar portfolio involving higher order terms: $\Sigma^{-1}[a(0) + \sigma a^{'}(0) + \sigma^2 a''(0)/2]$. Note that the equilibrium expressions of the slope $a^{'}(0)$ and curvature $a''(0)$ of the vector of risk premiums involves some additional mutual funds which will be further discussed in the next subsection.

Second, the beliefs-distorted portfolio $\varsigma_b^s = \Sigma^{-1}[\alpha_{\bullet s}\odot a(0)]$ is now completed by a higher order isomorphic term $\varsigma_b^{s\prime} = \Sigma^{-1}[\alpha_{\bullet s}\odot a'(0)]$. Note again that the coefficients of these beliefs-based portfolios in the investor s demand do not involve skewness, kurtosis or preferences for them.

Third, note that the skewness portfolio $\varsigma_{sk}= \Sigma^{-1}c$ is now augmented with a kurtosis portfolio $\varsigma_{kurt} = \Sigma^{-1}d$. Its coefficients in the asset demand of investor do not involve beliefs distortions and are simply proportional to respectively the intensities ρ_s and κ_s of investor s skew and kurtosis tolerances. The coefficients c_i and d_i measure respectively the contributions of asset i in the skewness and kurtosis of the mean-variance portfolio ς. In particular, up to correlation effects, a large co-skewness c_i increases the demand for asset i in proportion of investor s preference for positive skewness ρ_s, a large co-kurtosis d_i will decrease his/her demand for asset i in proportion of his/her aversion for kurtosis κ_s. Therefore, individual preferences for large positive skewness (resp. small kurtosis) will increase (resp. decrease), ceteris paribus, the equilibrium price of assets with positive co-skewness (resp. positive co-kurtosis). These effects will appear in the equilibrium values $a'(0)$ and $a''(0)$ of risk premiums slopes and curvatures.

Recall that Chabi-Yo, Leisen, and Renault (2006) showed that the affine regression of the squared return $(\varsigma^\intercal\mathbf{R})^2$ of the mean-variance mutual fund on the vector $\mathbf{R} = (\mathbf{R}_i)_{1\le i\le n}$ of returns on risky assets is an affine function of the skewness portfolio ς_{sk}. This yielded an alternative interpretation of the skewness portfolio. Not surprisingly, following similar arguments, one can show that the affine regression on $\mathbf{R}$ of the cubic return $(\varsigma^\intercal\mathbf{R})^3$ is an affine function of the kurtosis portfolio ς_{kurt}. Hence, the skewness and the kurtosis portfolios are respectively the best mimicking portfolios for the nonlinear payoffs $(\varsigma^{\intercal}\mathbf{R})^2$ and $(\varsigma^{\intercal}\mathbf{R})^3$.

As explained in section 2, in the case of joint normality of returns, the skewness portfolio ς_{sk} vanishes while the kurtosis portfolio ς_{kurt} is simply proportional to the mean-variance portfolio ς.

Finally, beyond the skewness portfolio $\varsigma_{sk} = \Sigma^{-1}c$ built on the vector of co-skewnesses $c_i = \varsigma^{\intercal}\Gamma_i\varsigma$, several portfolios are associated to various cross-co-skewnesses. We previously considered a cross-co-skewness measure between the mean-variance mutual fund ς and the beliefs portfolio ς_b :

$$c_{ib}^s = \varsigma^{\intercal}\Gamma_i\varsigma_b^s = Cov[(\varsigma^{\intercal}Y)((\varsigma_b^s)^{\intercal}Y), Y_i] \tag{28}$$

Along similar lines, we also define a cross-co-skewness measure between the mean-variance mutual fund ς and the skewness portfolio ς_{sk} as well as a cross-co-skewness measure with the "differentiated "portfolio $\varsigma' = \Sigma^{-1}a(0)$:

$$c_{isk} = \varsigma^{\intercal}\Gamma_i\varsigma_{sk} = Cov[(\varsigma^{\intercal}Y)(\varsigma_{sk}^{\intercal}Y), Y_i]$$

$$c_{i*} = \varsigma^{\intercal}\Gamma_i\varsigma' = Cov[(\varsigma^{\intercal}Y)(\varsigma'^{\intercal}Y), Y_i]$$

It is worth parenthetically recalling that all these cross-co-skewnesses are zero - like the co-skewnesses c_i - when the primitive asset return distributions are symmetric ($\Gamma_i = 0$). Moreover, Theorem 3.1 shows that cross-co-skewnesses give rise to additional portfolios, $\Sigma^{-1}c_b^s$, $\Sigma^{-1}c_{sk}$ and $\Sigma^{-1}c_*$ that gain importance in the portfolio decisions of investor s as his/her skewness tolerance ρ_s increases.

Finally we should note that a more illuminating interpretation of the various portfolios as mutual funds will emerge in the next subsection when we consider the quadratic approximation of the vector of equilibrium risk premiums.

3.2 Equilibrium Prices and Agent Demands

We need again to define some aggregate quantities in order to characterize equilibrium quantities. In particular, in addition to Definition 2.1, we have:

Definition 3.1 *In addition to the average investor characteristics in Definition 2.1 we also have the following population average with regards to kurtosis preferences:*

$$\overline{\kappa} = \sum_{s=1}^{S} \kappa_s \frac{\tau_s}{\sum \tau_s} \tag{29}$$

To avoid the proliferation of mutual funds, we simplify the structure of beliefs distortions by reinforcing Assumption 2.2 in the following way:

Assumption 3.1 *There is a one-dimensional structure of beliefs distortions. For all* $i = 1, \ldots, n$, *and* $s = 1, \ldots, S$:

$$h_{is} = \alpha_{is} a_i(0) \qquad h'_{is} = \alpha_{is} a_i(0)'$$

and $\alpha_{is} = \lambda_s \ \alpha_i$.

Note that Assumption 3.1 implies an even simpler framework compared to Assumption 2.2 with $J = 1$. To simplify notation, we simply write $\alpha_{is} = \lambda_s \ \alpha_i$ instead of $\alpha_{is} = \lambda_{0s} + \lambda_{1s}\alpha_i$. We use again the market clearing condition, namely $\sum_{s=1}^{S} q_s \omega''_s(0) = 0$ to derive the component $a''(0)$ of the equilibrium risk premiums, namely:

$$\begin{aligned} a''(0) &= \overline{\kappa} d - 2\overline{\lambda}(\alpha \odot a^{'}(0)) - 2(3\overline{\rho} - 1)(\varsigma^{\intercal}\Sigma\varsigma)a(0) \\ &\quad -4(\overline{\rho^2} - \overline{\rho}^2)c_{sk} - 4(\overline{\rho\lambda} - \overline{\lambda}\overline{\rho})c_b \end{aligned}$$

with $c_{sk} = (c_{isk})_{1 \leq i \leq n}$ with $c_{isk} = \varsigma^{\intercal}\Gamma_i\varsigma_{sk}$ and $c_b = (c_{ib})_{1 \leq i \leq n}$ with $c_{ib} = \varsigma^{\intercal}\Gamma_i\varsigma_b$ and average quantities defined as:

$$\overline{\rho^2} = \sum_{s=1}^{S} \rho^2 \frac{\tau_s}{\tau_s} \qquad \overline{\lambda\rho} = \sum_{s=1}^{S} \rho_s \lambda_s \frac{\tau_s}{\sum_{s=1}^{S} \tau_s} \tag{30}$$

Therefore, to summarize, we have established the following:

Theorem 3.2 *The vector of asset risk premiums in equilibrium,* $E\mathbf{R} - \mathbf{R}_f\mathbf{1} = \sigma^2 a(\sigma)$, *admits a second order Taylor expansion in the neighborhood of zero risk characterized by:*

$$\begin{aligned} a(0) &= \Sigma\overline{\omega}/\overline{\tau} \\ a^{'}(0) &= -\overline{\lambda}(\alpha \odot a(0)) - \overline{\rho}c \\ a''(0) &= \overline{\kappa}d - 2(3\overline{\rho} - 1)(\overline{\omega}^{\intercal}\Sigma\overline{\omega})\Sigma\overline{\omega}/\overline{\tau}^3 \\ &\quad -4(\overline{\rho^2} - \overline{\rho}^2)c_{sk} - 4(\overline{\rho\lambda} - \overline{\lambda}\overline{\rho})c_b \\ &\quad +2\overline{\lambda}\overline{\rho}(\alpha \odot c) + 2\overline{\lambda}^2[\alpha \odot (\alpha \odot a(0))] \end{aligned} \tag{31}$$

Proof: See Appendix E.

The difference with the risk premiums obtained in the context of a mean-variance-skewness investor is the term $a''(0)$ which can be decomposed in several components. The first three components would appear even without beliefs distortions. Hence, they are comparable to results in the literature on preferences for higher order moments where indeed a price $\overline{\kappa}d$ for low kurtosis is also found. The latter price is proportional to the average kurtosis aversion $\overline{\kappa}$ and the vector d of co-kurtosis coefficients $d_i = Cov[(\varsigma^{\intercal}Y)^3, Y_i]$. This term is similar to $\overline{\rho}c$ which determines the price for high skewness, notably the focus of interest in the cubic pricing kernel of Dittmar (2002). The main difference with Dittmar (2002) is that we do not operate within the representative agent paradigm. The standard approach to relaxing this paradigm (see e.g. Constantinides and Duffie (1996) and references therein) is to introduce both incomplete consumption insurance and consumption heterogeneity. While we also have investor heterogeneity, we have a different approach to market incompleteness. Recall that the role of the skewness portfolio ς_{sk} is to track the squared market return $(\varsigma^{\intercal}\mathbf{R})^2$, namely the affine regression of $(\varsigma^{\intercal}\mathbf{R})^2$ on $\mathbf{R}$ is an affine function of ς_{sk}. Therefore, in terms of quadratic hedging errors, the optimal way to hedge the risk $(\varsigma^{\intercal}\mathbf{R})^2$ with a portfolio comprising the risk-free rate and n risky assets $\mathbf{R}_i$, $i = 1, \ldots, n$, is to use the skewness portfolio ς_{sk}. However, markets are not complete with regards to this *"quadratic risk."* There is a non-zero residual risk, as in general:

$$Var[(\varsigma^{\intercal}\mathbf{R})^2] \geq Var[(\varsigma_{sk}^{\intercal}\mathbf{R})]$$

We argue that correctly taking this residual risk into account is what creates a wedge between our heterogeneous agent pricing model (without beliefs distortions) and the representative agent model used in Dittmar (2002). In the latter case, a Taylor expansion of the utility function one only finds a compensation for co-kurtosis coefficients d_i, while we find in addition to this a compensation for cross-skewness coefficients $c_{isk} = Cov[(\varsigma^{\intercal}Y)(\varsigma_{sk}^{\intercal}Y), Y_i]$. These coefficients measure the contribution of asset i to the aggregate risk:

$$\sum_{i=1}^{n} \varsigma_i c_{isk} = Cov[(\varsigma^{\intercal}Y)^2, (\varsigma_{sk}^{\intercal}Y)] = E[(\varsigma^{\intercal}Y)^2(\varsigma_{sk}^{\intercal}Y)]$$

In contrast, the co-kurtosis coefficients d_i measure the contribution of asset i to the market kurtosis:

$$\sum_{i=1}^{n} \varsigma_i d_i = Cov[(\varsigma^{\intercal}Y)^3, (\varsigma^{\intercal}Y)] = E[(\varsigma^{\intercal}Y)^4]$$

The difference between these two aggregate risks comes entirely from the differences between $(\varsigma^{\intercal}\mathbf{R})^2$, and what is hedged, namely $(\varsigma_{sk}^{\intercal}\mathbf{R})$.

It is interesting to note here that the approach we adopt, namely to combine investor heterogeneity and incomplete hedging, leads to conclusions strikingly similar to those of Constantinides and Duffie (1996). In their case, the pricing effects of heterogeneity are proportional to the cross-sectional variation of individual consumers' consumption growth. In our case, we find pricing effects that are proportional to the cross-sectional variance $(\overline{\rho^2} - \overline{\rho}^2)$ of individual investors' tolerance for skewness. Our approach has some clear advantages as far as empirical analysis is concerned as we treat the determinants of the cross-sectional variance as structural parameters that can be estimated using asset pricing time series data. Instead, the empirical implementation of Constantinides and Duffie (1996) requires, as they note, individual consumption/portfolio choice panel data.

The result in Theorem 3.2 adds two or possibly three priced factors, in addition those due to heterogeneous skewness preferences. To facilitate the interpretation of the additional factors, we will assume that:

Assumption 3.2 *All the non-zero coefficients of α are equal and therefore:*

$$\alpha \odot a(0) \propto \alpha \odot (\alpha \odot a(0))$$

meaning that both are proportional.

Assumption 3.2 will also be used in the empirical analysis reported in section 4. With Assumption 3.2 only the assets i for which $\alpha_i \neq 0$, i.e. assets for which beliefs distortions matter, are selected. Their pricing effects, characterized by betas with respect to the beliefs portfolios $\varsigma_b = \Sigma^{-1}[\alpha \odot a(0)]$, are significant once there is a non-zero aggregate beliefs distortion $\overline{\lambda}$.

When both aggregate beliefs distortion $\overline{\lambda}$ and aggregate skew tolerance $\overline{\rho}$ are non-zero, an additional priced factor emerges, as shown in in Theorem 3.2. This factor is characterized by $\alpha \odot c$, namely, whenever $\alpha_i \neq 0$ and price effect of co-skewness appears through c_i. Abandoning the representative agent setting also adds a factor pertaining to the covariation of skewness preferences and beliefs distortions. That is, the non-zero cross-sectional covariance $\overline{\rho\lambda}$ - $\overline{\rho}$ $\overline{\lambda}$ yields a cross-co-skewness factor determined by:

$$c_{ib} = Cov((\varsigma^{\intercal})(\varsigma_b^{\intercal}), Y_i)$$

These coefficient represent the contribution of asset i to the covariance between the mean-variance portfolio and the skewness portfolio:

$$\sum_{i=1}^{n} \varsigma_i c_{ib} = Cov[(\varsigma^{\intercal})^2, (\varsigma_b^{\intercal})] = Cov[(\varsigma_{sk}^{\intercal}), (\varsigma_b^{\intercal})]$$

since the hedging error $(\varsigma^{\intercal})^2$ - $(\varsigma_{sk}^{\intercal})$ is uncorrelated with all linear portfolios. Hence, the cross-sectional covariance between skew tolerance and beliefs distortion factors has a pricing effect when the skewness and beliefs portfolios are correlated.

Given assets risk premiums in equilibrium, we rewrite agent's portfolio weights given in Theorem 3.1:

Theorem 3.3 *Assume the setting of Theorem 1.2, with preferences be specified as in (5), where ρ_s and κ_s denote respectively the skewness and kurtosis tolerances. Then, in the neighborhood of $\sigma = 0$, the second order approximation $[\omega_s(0) + \sigma\omega_s^{'}(0) + \sigma^2\omega_s^{''}(0)/2]$ of the vector $\omega_s(\sigma)$ of shares of wealth invested is defined by:*

$$q_s\omega_s(\sigma) = q_s[\omega_s(0) + \sigma\omega_s^{'}(0) + \sigma^2\omega_s^{''}(0)/2] \tag{32}$$

with:

$$\begin{array}{rcl} q_s\omega^s(0) & = & \tau_s(\overline{\omega}/\overline{\tau}) \\ q_s\omega_s^{'}(0) & = & \tau_s(\rho_s - \overline{\rho})\varsigma_{sk} + (\tau_s/\overline{\tau})(\lambda_s - \overline{\lambda})\varsigma_b \\ q_s\omega_s^{''}(0)/2 & = & \tau_s(\lambda_s - \overline{\lambda})\varsigma_b^{\prime} - \tau_s/2(\kappa_s - \overline{\kappa})\varsigma_{kurt} \\ & & 3(\tau_s/\overline{\tau}^3)(\rho_s - \overline{\rho})(\omega^{\intercal}\Sigma\overline{\omega})\overline{\omega} \\ & & +2\tau_s[\rho_s(\rho_s - \overline{\rho}) - (\overline{\rho^2} - \overline{\rho}^2)]\Sigma^{-1}c_{sk} \\ & & +2\tau_s[\rho_s(\lambda_s - \overline{\lambda}) - (\overline{\lambda\rho} - \overline{\rho}\overline{\lambda})]\Sigma^{-1}c_b \end{array}$$

where:

$$\begin{array}{rclrcl} \varsigma_{sk} & = & \Sigma^{-1}c & \varsigma_b & = & \sigma^{-1}[\alpha \odot \Sigma\overline{\omega}] \\ \varsigma_{kurt} & = & \Sigma^{-1}d & \varsigma^{\prime} & = & \Sigma^{-1}a^{\prime}(0) \\ \varsigma_b^{\prime} & = & \Sigma^{-1}[\alpha \odot a^{\prime}(0)] & & & \end{array}$$

where $a^{\prime}(0) = -\overline{\lambda}\alpha \odot a(0)$ - $\overline{\rho}c$ and $a(0) = \Sigma(\overline{\omega}/\overline{\tau})$, and for $i = 1, \ldots, n : c_{isk} = \varsigma^{\intercal}\Gamma_i\varsigma_{sk}$, c_{ib} $= \varsigma^{\intercal}\Gamma_i\varsigma_b^{\prime}$.

Proof: See Appendix F.

To summarize, equilibrium individual asset demands $\omega_s(\sigma)$, as represented by their second order expansion as appearing in equation (32), involves the following mutual funds in addition to the mean-variance portfolio $\varsigma = \overline{\omega}/\overline{\tau}$:

- the beliefs portfolio $\varsigma_b = \Sigma^{-1}[\alpha \odot \Sigma\overline{\omega}]$
- the skewness portfolio $\varsigma_{sk} = \Sigma^{-1}c$, with $c_i = Cov[(\varsigma^{\intercal}Y)^2, Y_i]$
- the kurtosis portfolio $\varsigma_{kurt} = \Sigma^{-1}d$, with $d_i = Cov[(\varsigma^{\intercal}Y)^3, Y_i]$
- the cross-co-skewness portfolio: $\varsigma_{csk} = \Sigma^{-1}c_{sk}$, with $c_{isk} = Cov[(\varsigma^{\intercal}Y)(\varsigma_{sk}^{\intercal}Y), Y_i]$
- the beliefs-about-skewness portfolio $\varsigma_{bc} = \Sigma^{-1}(\alpha \odot c)$
- the cross-co-skewness beliefs portfolio $\varsigma_{bsk} = \Sigma^{-1}c_b$, with $c_{ib} = Cov[(\varsigma^{\intercal}Y)(\varsigma_b^{\intercal}Y), Y_i]$

A particular investor s will hold in equilibrium shares of some mutual funds, where the shares depend on the spread between his/her preference/beliefs characteristics and the economy-wide averages:

Mutual funds	Corresponding Preference/Beliefs Characteristics
Beliefs portfolio	λ_s
Skewness portfolio	ρ_s
Kurtosis portfolio	κ_s
Cross-co-skewness portfolio	$\rho_s(\rho_s - \overline{\rho})$
Beliefs-about-skewness portfolio	λ_s
Cross-co-skewness beliefs portfolio	$\rho_s(\lambda_s - \overline{\lambda})$

3.3 The Pricing Kernel

In equilibrium, the pricing kernel of the average investor should correctly price the vector of returns as shown in expression (25).

As in section 2.3, we determine the factors which span the pricing kernel, using the mutual funds that appear in the equilibrium asset demands. Therefore, when the risk premium vector is approximated by its second order expansion $[a(0) + \sigma a'(0) + (\sigma^2/2)a''(0)]$, we expect the pricing kernel to be spanned by:

- First, the factors appearing in Theorem 3.2 for the first approximation of the risk premium vector, namely:
  - (1) the risk-free rate $\mathbf{R}_f$

- (2) the market return $\mathbf{R}_M = \overline{\omega}^\intercal \mathbf{R}$
- (3) the payoff on the beliefs portfolio $\mathbf{R}_{Mb} = \overline{\tau}\varsigma^\intercal \mathbf{R}$
- (4) the quadratic market return $\mathbf{R}_M^2$

- Second, four additional factors corresponding to the supplementary mutual funds appearing in Theorem 3.3:
  - (5) the cubic market return $\mathbf{R}_M^3$
  - (6) the product of the market return and the payoff on the skewness portfolio
  - (7) the product of the market return and the payoff on the beliefs portfolio
  - (8) a beliefs-about-skewness portfolio

The cubic market return characterizes the reward on the kurtosis portfolio which can be interpreted as the affine regression of $\mathbf{R}_M^3$ on the vector $\mathbf{R}$ of primitive returns. The product of the market return and the payoff on the skewness (respectively beliefs) portfolio, characterizes the reward on the cross-co-skewness (respectively cross-co-skewness beliefs) portfolio. Finally, while the beliefs portfolio $\varsigma_b = \Sigma^{-1}[\alpha \odot a(0)]$ (and its payoff $\mathbf{R}_{Mb} = \overline{\tau}\varsigma^\intercal \mathbf{R}$) was introduced by the beliefs about the vector $a(0)$ of risk premiums, we have also to introduce a beliefs-about-skewness portfolio with payoff $\mathbf{R}_{bc} = \overline{\tau}\varsigma_{bc}^\intercal \mathbf{R} = \overline{\tau}(\alpha \odot c)^\intercal \Sigma^{-1}\mathbf{R}$. For the sake of simplicity, we define the random variables:

$$\begin{aligned}
md(\mathbf{R}_{M,skew}) &= Var(\mathbf{R})^{-1}Cov((md(\mathbf{R}_M))^2, \mathbf{R})md(\mathbf{R}) \\
md(\mathbf{R}_{Mb}^*) &= (Var(\mathbf{R})^{-1}Cov(\mathbf{R}_{Mb}, \mathbf{R}))md(\mathbf{R}) \\
md(\mathbf{R}_{Mb}^o) &= (Var(\mathbf{R}))^{-1}((\alpha \odot \alpha) \odot Cov(\mathbf{R}_M, \mathbf{R}))md(\mathbf{R}) \\
md(\mathbf{R}_{Mbskew}^o) &= (Var(\mathbf{R}))^{-1}(\alpha \odot Cov((md(\mathbf{R}_M))^2, \mathbf{R}))md(\mathbf{R})
\end{aligned}$$

where $md(X)$ denotes the mean-deviation part of a random variable X, and $md([md(X)]^k)$ corresponds to $(X - E(X))^k$ - $E(X - E(X))^k$. Then the following theorem holds:

Theorem 3.4 *In equilibrium, assuming Assumptions 3.1 and 3.2 to hold, then the vector of assets risk premiums*

$$\sigma^2 a(\sigma) = \sigma^2[a(0) + \sigma a'(0) + \sigma^2 a''(0)/2]$$

is consistent with the pricing kernel:

$$\begin{aligned}
m = {} & A_0 + A_1 md(\mathbf{R}_M) + A_2 md(\mathbf{R}_{Mb}) + A_3 md((md(\mathbf{R}_M))^2) + A_4 md((md(\mathbf{R}_M))^3) \\
& + A_5 md(md(\mathbf{R}_M)md(\mathbf{R}_{M,skew})) + A_6 md(md(\mathbf{R}_M)md(\mathbf{R}_{Mb}^*)) \\
& + A_7 md(\mathbf{R}_{Mb}^o) + A_8 md(\mathbf{R}_{Mbskew}^o)
\end{aligned}$$

with:

$$\begin{aligned}
A_0 &= 1/\mathbf{R}_f & A_1 &= -1/(\overline{\tau}\mathbf{R}_f) + 3(\overline{\rho} - 1)Var(\mathbf{R}_M)/(\overline{\tau}^3\mathbf{R}_f) \\
A_2 &= \overline{\sigma\lambda}/(\mathbf{R}_f\overline{\tau}) & A_3 &= \overline{\rho}/\mathbf{R}_f\overline{\tau}^2) \\
A_4 &= -\overline{\kappa}/2\mathbf{R}_f\overline{\tau}^3) & A_5 &= 2(\overline{\rho^2} - \overline{\rho}^2)/(\overline{\tau}^3\mathbf{R}_f) \\
A_6 &= 2(\overline{\rho\sigma\lambda} - \overline{\sigma\lambda}\overline{\rho})/(\overline{\tau}^2\mathbf{R}_f) & A_7 &= -\left(\overline{\sigma\lambda}\right)^2(\overline{\tau}\mathbf{R}_f) \\
A_8 &= -\overline{\rho}\left(\overline{\sigma\lambda}\right)/\mathbf{R}_f\overline{\tau}^2)
\end{aligned}$$

where $md(X)$ denotes the mean-deviation part of a random variable X, and $md([md(X)]^k)$ corresponds to $(X - E(X))^k$ - $E(X - E(X))^k$.

Proof: See Appendix G.

Note that with Assumption 3.2, we can regroup the terms $A_2 md(\mathbf{R}_{Mb})$ and $A_7 md(\mathbf{R}_{Mbc})$, which combined yields a single factor defined by the beliefs portfolio $\mathbf{R}_{Mb}$. This pricing kernel generalizes the market co-skewness and co-kurtosis models that are analyzed by Harvey and Siddique (2000), Dittmar (2002) and many others. Distortion and heterogeneity of beliefs introduce additional factors in the pricing kernel that were defined above. If there are no beliefs distortions, then $A_2 = A_6 = A_7 = A_8 = 0$. In this case, the pricing kernel has the same functional form as the one proposed by Dittmar (2002), with the exception that there is an additional factor which captures the dispersion of investors skew tolerance parameters. This factor will not show up if one uses Taylor expansion series as in Dittmar (2002). An interesting exercise is to investigate whether additional factors that appear in the pricing kernel are economically relevant.

4 Empirical Analysis

In this section we assess the empirical evidence pertaining to the pricing kernels for the mean-variance-skewness and mean-variance-skewness-kurtosis asset pricing models. We need to augment the notation to present the empirical analysis. First, we collect all the (structural) parameters of a model into a parameter vector θ. Obviously, depending on the model, the number of parameters will differ. The specification of the parameter vector will be left general for the moment until we discuss specific models. Moreover, so far none of the variables had time indices - they will from now on. For example, the pricing kernel was denoted by m. For the purpose of empirical analysis we will denote it by $m_{t+1}(\theta)$, indicating its dependence on the value of the parameter vector and its dependence on time. Finally, the empirical implementation will need to make some simplifications in order to keep the number of factors, and their interpretation tractable. Most importantly, we will assume Assumptions 3.1 and 3.2 to hold. The former reduces the beliefs portfolio to a single one, which corresponds to the setup in Anderson, Ghysels, and Juergens (2005a).

In its generic form, the fundamental pricing equation can then be written as

$$E\mathbf{R}_{t+1} m_{t+1}(\theta) = \mathbf{1}$$

where, as noted before, $\mathbf{1}$ is the sum vector $(1, 1, ..., 1)^{\intercal}$ of dimension n, $\mathbf{R}_{t+1}$ is a vector of returns on n risky assets. We write the vector of errors as follows:

$$\varepsilon_{t+1} = (\mathbf{R}_{t+1}) m_{t+1}(\theta) - \mathbf{1} \tag{33}$$

and denote the sample counterpart of $E\varepsilon_{t+1}$ as:

$$g_T(\theta) = \frac{1}{T} \sum_{t=1}^{T} \varepsilon_{t+1} \equiv E_T[\varepsilon_{t+1}] \tag{34}$$

where T represents the number of time series observations and $E_T[.]$ is a sample average operator. A test of model specification can be obtained by minimizing the quadratic form:

$$J = g_T(\theta)' W_T g_T(\theta)$$

In the estimation process, we use the iterated GMM and the Hansen and Jagannathan (1997) method. The latter specifies the weighting matrix W_T as the second moment of instrument-scaled returns:

$$W_T = (E_T \mathbf{R}_{t+1} \mathbf{R}'_{t+1})^{-1}$$

To estimate all the models, we utilize the returns on 20 industry sorted portfolios, where the industry definitions follow the two-digit SIC codes used in Moskowitz and Grinblatt (1990) and are described in Table 1. Our empirical analysis is most closely related to that of Dittmar (2002) who also uses 20 industry portfolios. This will allow us to draw some comparisons between our empirical results and those of Dittmar (2002). The data is monthly and spans the period December 1981 until December 2005, or a total of 289 monthly returns. As proxy for the market return, $\mathbf{R}_M$, we use the CRSP value-weighted index. To proxy the beliefs market return $\mathbf{R}_{Mb}$ we utilize the dispersion returns used in Anderson, Ghysels, and Juergens (2005a), which consists of portfolios of stocks with strong disagreement among analysts. In order to construct the portfolios based on exposure to analyst dispersion, we utilize the Thomson Financial/IBES Summary History file. The Summary History file contains distributional data on earnings forecasts by investment analysts, including consensus and median forecasts, standard deviation, high and low forecasts, and the number of analysts providing forecasts. TF/IBES updates the Summary History file on the third Wednesday of each month. The variable of interest for our purposes is the standard deviation, which measures the dispersion of analysts' beliefs in their forecasts. We use two observational periods in order to measure dispersion: the one-year ahead annual earnings (short-term) forecast and the five-year earnings growth (long-term) forecast. In each month, we identify all firms that have at least three analysts providing forecasts for each stock in order to get meaningful dispersion measures. For robustness, we also require the number of analysts at five and ten as well. In our empirical analysis we report the case of 3 and 10. In order to construct high and low exposure portfolios, we rank firms according to their dispersion of analyst forecasts (short-term and long-term individually) in each month. The first method employed is to rank stocks above the median as high dispersion stocks while those below the median are low dispersion stocks. In the second method, stocks are separated into deciles, to capture the extremes, only the high dispersion stocks (decile 10) is considered. Stocks are value-weighted in each portfolio.

The empirical analysis will involve various approaches. First we will appraise theoretical pricing kernel against Hansen-Jagannathan volatility bounds. Next, we will estimate the structural parameters via GMM. A third and final empirical approach will be related to betas and market prices of risk. In particular, the mean-variance-skewness and -kurtosis pricing kernel can be rewritten as:

$$m(\theta) = A^{\intercal} F \tag{35}$$

where A is a vector with component A_i. For each industry portfolio, the pricing relation $Em\mathbf{R}_i = 1$ implies the beta-pricing relationship:

$$E\mathbf{R}_i - \mathbf{R}_f = -\mathbf{R}_f Cov(m, \mathbf{R}_i) = \beta_i \Lambda = \sum_{j=1}^{K} \beta_{ij} \Lambda_j \tag{36}$$

where $\beta_i = Cov(\mathbf{R}_i, F)(Var(F))^{-1}$ and $\Lambda = -\mathbf{R}_f Var(F)A$. The vector Λ is the market price of risk.

The remainder of this section is structured as follows. In a first subsection, we start with assessing the empirical relevance of our models using Hansen-Jagannathan bounds. Next, we estimate the structural parameters for the mean-variance-skewness and -kurtosis models. In addition we also compute the implied nonstructural pricing kernel parameters. The final subsection is devoted to betas and prices of risk.

4.1 Hansen and Jagannathan Bounds

To be included

4.2 Estimation of the Mean-Variance-Skewness Pricing Kernel

The empirical results for the pricing kernel with skewness appear in Tables 3 and 4. The former contains the structural parameter estimates. In Table 3 the coefficients are estimated using an iterated GMM approach and using the Hansen and Jagannathan (1997) weighting matrix. We utilize the returns on 20 industry sorted-sorted portfolios appearing in Table 1. As proxy for the market return, $\mathbf{R}_M$, we use the CRSP value-weighted index. To proxy the "beliefs market return" we utilize the dispersion returns used in Anderson, Ghysels, and Juergens (2005a), which consists of portfolios of stock with large disagreement among analysts. Two portfolios are considered, long-run dispersion denoted $HighLTD$ - Top Decile and short-run forecast dispersion denoted $HighSTD$ - Top Decile. Dispersion portfolios are reported with at least 10 and at least 3 analysts coverages respectively. Hence, the parameters are estimated for four beliefs portfolio returns, respectively using GMM and the Hansen and Jagannathan (1997) weighting matrix. For each model specification, the first line represents the parameter estimates. The second line represents the p-value of the parameter estimates. The third line contains the objective function value, the p-value for the model fit and the J-statistics.

First of all, we find that the overall model fit is remarkably good, as the J-statistics yield non-rejections of the overidentifying restrictions. We have a total of 21 moment conditions and four structural parameters, hence 17 overidentifying restrictions, which is typical for many empirical asset pricing models. The overall empirical success of these models is therefore remarkable. In addition, we also find reasonable values for the implied risk parameters. Namely, if we interpret the risk aversion tolerance in terms of a CRRA utility function, as in equation (**??**), we find risk aversion levels ranging from 7.9 to 21.3 (see the columns $1/\overline{\tau}$). We also find reasonable levels of implied risk-free rates, particularly with the Hansen and Jagannathan (1997) weighting matrix (see columns $1/\mathbf{R}_f$ with values ranging from 1.05 to .995).

The skewness parameter $\overline{\rho}$ estimates vary between .08 and .69. While the parameters have the right sign and magnitude, none of the estimates are statistically significant. In contrast, for the heterogeneity of beliefs parameter $\overline{\sigma\lambda}$ we find strongly statistically significant values ranging from .37 to .87.

To appraise the market price of risk implications of the structural parameters we also report in Table 4 the coefficients A_0, A_1, A_2, A_3 implied by the estimated structural parameters reported in Table 3. We find that all four parameters are statistically significant (although among the four A_3 is somewhat borderline). Recall that our empirical analysis is most closely related to that of Dittmar (2002) who also uses 20 industry portfolios. Before we draw comparisons, it should be noted that Dittmar includes a human capital factor into his specification of the pricing kernel, while in our specification we include a beliefs portfolio. Table III of Dittmar (2002) presents results of specification tests when the measure of aggregate wealth does not include human capital. These results are therefore the closest to our specification (ignoring the beliefs portfolio) when estimated with the H-J weighting matrix also used by Dittmar. Apart from these difference, it should be noted that Dittmar consider a pricing kernel with time-varying coefficients. Table III in Dittmar (2002) contains the sample average values of the coefficients. Our estimates of the intercept, i.e. A_0 in Table 4, roughly .99 with the H-J weighting, are in the same range (between 0.94 and 1.08) as the corresponding estimates of $d(\bar{Z})_{0t}$ in Table III of Dittmar (2002). Similarly, Dittmar's estimate of $d(\bar{Z})_{2t}$ (equal to 77.6) are directly comparable with our H-J weighting estimates of A_3 (ranging from 40.82 and 84.33).

Overall for the mean-variance-skewness model we find that, all structural parameters have the expected sign. It seems that beliefs are more important than skewness because all beliefs parameters are statistically significant while skewness parameters are not (given our sample size). However there are two reason to be cautious. Table 4 indicates that the (skewness) non-structural parameter A_3 is significant - although obviously since $A_3 = \overline{\rho}/(\mathbf{R}_f\overline{\tau}^2)$ in this model the statistical significance of the reduced form parameter is not only due to the skewness parameter $\overline{\rho}$. Moreover, as shown in Table 9, the market price of risk indicates that the skewness risk is significant and it is priced. The market price of risk also has the expected sign.

4.3 Estimation of the Mean-Variance-Skewness-Kurtosis Pricing Kernel

The empirical results for the pricing kernel with skewness and kurtosis preferences appear in Tables 5 and 6. The former contains the structural parameter estimates. The top panel shows the parameters when the iterated GMM approach is used while the bottom panel presents the parameters estimated with the Hansen and Jagannathan (1997) weighting matrix. We utilize again the returns on 20 industry sorted-portfolios appearing in Table 1, and the market return and beliefs portfolio returns described in the previous section. The first line presents the parameter estimates and the second line presents the p-values. The third line contains the objective function value, the p-values for the model fit and the J-statistics.

First of all, the J-statistics show that the mean-variance-skewness-kurtosis pricing kernel fit is slightly better than the mean-variance-skewness pricing kernel fit. The distance measure and the p-values for the pricing kernels suggest significant improvements moving from a pricing kernel with skewness (see Table 3) to the pricing kernel with skewness and kurtosis preferences. The implied risk aversion parameters ($1/\overline{\tau}$) are reasonable. When High STD return is used to proxy the "market beliefs return", the risk aversion parameters are all statistically significant except when the Hansen and Jagannathan weighting matrix and High STD return with 10 and more analysts are used. In addition, when the Hansen and Jagannathan weighting matrix is used, the risk aversion parameter is lower than the one obtained with the mean-variance-skewness pricing kernel. The risk aversion levels range from 3.76 to 20.99. We also find reasonable level of the risk-free rate, particularly with the Hansen and Jagannathan weighting matrix, with $E(m) = 1/\mathbf{R}_f$ roughly .995.

The estimates of the skewness parameter $\overline{\rho}$ vary between 0.15 and 0.50. While these parameters are positive as expected, there are significant at 1% level when the iterated GMM approach is used and the High STD return is a proxy for the "market beliefs return". The heterogeneity of beliefs parameter $\overline{\sigma\lambda}$ also has the expected sign. Furthermore, the estimates are statistically significant at 10% level when High STD return is used. Their values range from 0.60 to 1.26. The kurtosis tolerance parameter $\overline{\kappa}$ estimates vary between 0.10 and 23.37. While these parameters have the right sign and magnitude, none are statistically significant. The parameter $\overline{\rho^2}$ which captures the dispersion of investors' skew tolerance parameters has the right sign and ranges from 0.21 to 14.13. This parameter is significant only when the iterated GMM approach is used and when the High STD return with 3 and more analyst is used to proxy the "market beliefs return".

The parameter $\overline{\rho\sigma\lambda}$ which captures the interaction between the skew-tolerance parameter $\overline{\rho}$ and the beliefs parameter $\overline{\sigma\lambda}$ is positive and ranges from 0.37 to 1.86. When High STD return with 10 and more analysts is used to proxy the "beliefs market return", this parameter is statistically significant at 10% level and significant at 5% level with High STD return with 3 and more analysts as a proxy.

The estimate parameters displayed in Table 5 indicates that (see for example the results for High STD top Decile), the risk aversion $\frac{1}{\overline{\tau}}$, the average skew tolerance, $\overline{\rho}$, the parameter that capture dispersion of investor skew tolerance parameters, $\overline{\rho^2}$, and the parameters that captures interaction (covariance) of skewness preference and beliefs $\overline{\sigma\lambda\rho}$ are significant and have expected sign. The average kurtosis preference parameter $\overline{\kappa}$ is not significant.

To appraise the market price of risk implications of the structural parameters, we report in Table 6, the coefficients, A_0 through A_6 implied by the structural parameters reported in Table 6. These coefficients have the expected sign. Among those parameters only a few (A_0,

A_1, A_2) are statistically significant. While the parameters A_3, A_4, A_5, A_6 are not significant, the market price of risk implied by the corresponding factor loadings are statistically significant (see subsection 4.4 below). It does suggests that the loading factors are not independent. Moreover, there might also be time variation in parameters A_0 throughA_6, as say in Dittmar (2002), although such a specification does not naturally flow from our theoretical development.

4.4 Betas and Market Prices of Risk

In this final empirical subsection we turn our attention to market betas and price of risk. In Tables 7 and 8 we report the betas for the Mean-Variance-Skewness-Skewness and -Kurtosis models for the 20 industry portfolios. To interpret the results we also compute market prices of risk. These are reported in Tables 9 and 10.

They are as many betas as components to the pricing kernel, and therefore three betas for the Mean-Variance-Skewness model and seven betas for the Mean-Variance-Skewness-Kurtosis model. Hence in Table 8, which covers the former, we report three betas for the 20 industry sorted portfolio appearing in Table 1. Betas are computed for the full sample with respect to the excess market portfolio $r_M = \mathbf{R}_M - E\mathbf{R}_M$, (denoted β_{i2}), the excess "market beliefs return" $r_{Mb} = \mathbf{R}_{Mb} - E\mathbf{R}_{Mb}$, (denoted β_{i3}) and excess market squared return $\mathbf{R}_M^2 - E\mathbf{R}_M^2$, (denoted β_{i4}). Likewise in Table 8, betas are computed for the full sample with respect to the excess market portfolio $\mathbf{R}_M$, the excess "market beliefs return" $r_{Mb} = \mathbf{R}_{Mb} - E\mathbf{R}_{Mb}$, (denoted β_{i3}), the excess market squared return $r_M^2 - Er_M^2$, (denoted β_{i4}), the excess cubic market return $r_M^3 - Er_M^3$, (denoted β_{i5}), the cross-product of the market return r_M and a skewness portfolio excess return $r_{M,skew} = R_{M,skew} - R_{M,skew}$, (denoted β_{i6}), and finally the product of the market excess return r_M and the excess "market beliefs return" r_{Mb}.

A first subsection is devoted to the Mean-Variance-Skewness model, and a second to the model that includes kurtosis.

4.4.1 Mean-variance-Skewness Models

We will focus most of our discussion on the case *10 and more analyst-High LTD return*, the first panel in Table 7. This panel is representative of the results for the other cases. We note first that the coefficients β_{i2} are all positive. The lowest value 0.3484 is for Electrical Equipment while the highest value 0.7366 is for Primary Metals. For the sake of robustness, it is worth noting that for example in the case of *10 and more analyst High STD return* the lowest value is 0.1207 for Mining while the highest value 0.6816 for Electrical Equipment. Hence, the range is roughly similar.

To interpret further the case *10 and more analyst-High LTD return*, let us compute the market price of risk, assuming for instance the Iterated GMM approach is used to estimate the model (see Table 9 - the results are quite similar with the Hansen and Jagannathan weighting matrix). As shown in Table 9, the market prices of risk have the expected sign and are reasonable. For example, if the excess return on the market portfolio increases by 1 %, then the expected excess return on the Electrical Equipment industry portfolio increases by 0.64%, given it has a $\beta_{i2} = 0.3484$. Likewise, for the highest beta industry - Primary Metals - the expected excess return on Primary Metals increases by 1.36%.[5]

Next we turn to the beliefs portfolio $\mathbf{R}_{Mb}$, measured as *10 and more analyst-High LTD return,* and its associated beta β_{i3}. All are positive except for Utilities (-0.0162) and β_{i3} ranges from 0.01 to 0.68 (excluding the negative one). The lowest is for Food & Beverage industry portfolio and the highest for Electrical Equipment. Therefore, for the latter portfolio, if the return on the beliefs portfolio increases by 1 %, then the expected excess return on Electrical Equipment increases by 0.68 %. Note that this positive impact on returns of increased dispersion has been a bit controversial. Diether, Malloy, and Scherbina (2002) argued that increased dispersion led to decline in returns, whereas Anderson, Ghysels, and Juergens (2005a) found the opposite. The results reported here, obtained via a new model specification and a large set of industry portfolios seems to confirm that increases of dispersion of beliefs have a positive impact on returns as reported in Anderson, Ghysels, and Juergens (2005a). This result is also robust to changes in the beliefs portfolio. In the case of *10 and more analyst High STD return* the β_{i3} vary from 0.2697 to 0.6901. Hence, the results are even stronger. The highest beta is obtained for Primary Metals. This also applies to the *3 and more analyst High LTD and High STD returns* cases as well.

Finally, we turn to the skewness portfolio and its impact on expected excess returns as measured by the coefficients β_{i4}. They are negative for all industries and range from -2.5985 to -0.18. The highest value is observed for Textile Products. Since the market prices of risk of the skewness factor is also negative, an increase in the return on the skewness portfolio by 1 % will increases the expected excess return on Textile Products by 0.55%.

When we look to the averages across industries, he price of risk of the market return factor is significant and is positive. In addition, if the return on the market in excess of the risk-free rate increases by 1%, the excess average industry return whose beta is 0.5099, increases by 0.5%. We also find that the price of risk of the beliefs portfolio is significant and positive. If the excess return on the beliefs portfolio increases by 1% the excess average industry return whose beta is 0.1863, increases by 0.19%. The price of risk of the squared market return is significant and is negative. If the excess return on the squared market return increases by 1% the excess average industry return whose beta is -1.6683, decreases by 1.67%.

Continuing with industry averages, let us consider some of the other various on the beliefs portfolios. The finding that the price of risk of the market return factor is significant and is positive is robust. In addition, if the excess return on the market increases by 1%, the excess average industry return increases by 0.36% (High STD return with 10 or more analysts) 0.44% (High LTD return with 3 or more analysts), and finally 0.26% (High STD return with 3 or more analysts). All of this are results roughly of the same order of magnitude. We

[5] As noted before, similar results are obtained with *10 and more analyst High STD return.* For example, expected excess return on the Mining portfolio increases by 1.41 %.

also find that the price of risk of the beliefs portfolio is positive. If the risk premium on the beliefs portfolio return increases by 1% the excess average industry increases by 0.46% (High LTD return with 3 or more analysts), 0.24% (High LTD return with 3 or more analysts), and finally 0.57% (High STD return with 3 or more analysts). Note that this impacts are again roughly similar and of the same order of magnitude as the impact of excess returns on the market. Finally, the finding that the price of risk of the squared market return is significant and is negative is also robust for industry averages using other beliefs portfolio specifications. Notably, when the excess return on the squared market increases by 1% the excess average industry return decreases by 1.02% (High STD return with 10 or more analysts), 1.65% (High LTD return with 3 or more analysts), and finally 1.1% (High STD return with 3 or more analysts) .

4.4.2 Mean-variance-Skewness-Kurtosis Models

Let us start again with the *10 and more analyst High LTD return* case. As shown in Table 8, the coefficients β_{i2} are all positive and vary from 0.35 to 0.7385. The highest value is observed for Primary Metals. The average value of β_{i2} is 0.4780 with a standard deviation of 0.0990. Assume again that the Iterated GMM approach is used to estimate the model and compute the market prices of risk (see Table 10). As shown in Table 10, the market price of risk have expected sign and are quite reasonable. For Primary Metals, the market portfolio beta β_{i2} and price of risk imply an increase of the Primary Metals portfolio excess expected return by 1.30% upon a 1 % increase in the market return. On average, the increase of expected excess returns on industry portfolios increases by 0.85%.

When we consider industry averages and we find that the price of risk of the excess market return is significant and is positive. An increase of 1% moves the excess average industry return upward by 0.48% (High LTD return with 10 or more analysts) and 0.4% with all three other beliefs portfolio specifications. If the excess return on the beliefs portfolio return increases by 1% the excess average industry return increases by 0.19% (High LTD return with 10 or more analysts), 0.45% (High STD return with 10 or more analysts), 0.24% (High LTD return with 3 or more analysts), and finally 0.56% (High STD return with 3 or more analysts). Hence the impact is more dispersed but still of the same magnitude as the market portfolio. The price of risk of the squared market return is again significant and is negative and increases of its excess return by 1% impact the excess average industry negatively by 0.16% (High LTD return with 10 or more analysts), 0.7% (High STD return with 10 or more analysts), 0.2% (High LTD return with 3 or more analysts), and finally 0.3% (High STD return with 3 or more analysts). Note, however, that the betas are not significant. Hence, skewness seems not to significantly affect expected returns. Next we move to the price of risk of the cubic market return which is significant and is positive with premiums on the cubic market return increases by 0.1% resulting in the excess average industry return rising 0.46% (High LTD return with 10 or more analysts), 0.41% (High STD return with 10 or more analysts), 0.6% (High LTD return with 3 or more analysts), and finally 0.55% (High STD return with 3 or more analysts). This is a fairly stable pattern across different configurations, unlike the skewness case. The price of risk of the payoff defined by the cross product of the market return and the skewness payoff $R_{M,skew}$ is mostly positive and significant with the risk premium on this payoff increases by 1% resulting in rather small impacts on the average industry portfolios in the order of 0.1%. The last case we consider is the price of risk of the payoff defined by the cross product of the market return and the beliefs mutual fund R^*_{Mb}. An increase of excess return of this portfolio results in a negative and significant response of average industry returns of between 0.6 % and 0.8%.

5 Conclusions

In this paper we provide economic foundations, in terms of risk preferences and heterogenous beliefs, for pricing kernels that depend on higher order empirical moments of the return on the market portfolio. This far, the parameter estimates obtained by Bansal, Hsieh, and Viswanathan (1993), Chapman (1997), Dittmar (2002) and Harvey and Siddique (2000), among others, can be viewed as reduced form estimates to which we attach an economic interpretation in terms of attitudes towards risk and aggregation of heterogenous preferences.

The structural interpretation of the pricing kernel we obtain allowed us to disentangle the effects of heterogeneous beliefs and preferences on asset prices. Various additional pricing factors appear in the pricing kernel with weights depending on the dispersion across investors of preferences, beliefs and interactions between them. We also established a well-founded aggregation theory which tells us what are the relevant aggregates when investors are heterogeneous with respect to both preferences and beliefs. This is in contrast to the results of Gorman (1953). To handle the various sources of heterogeneity of beliefs and preferences, we follow the framework of Samuelson (1970) and its recent generalization of Chabi-Yo, Leisen, and Renault (2006). This generic approach allowed us to derive, for risks that are infinitely small, optimal shares of wealth invested in each security coincide with those of a Mean-Variance-Skewness-Kurtosis optimizing agent. Through these local approximations we are able to tease out the various sources of risk.

We show that heterogeneity of beliefs as specified by only appears at higher orders. Consequently, heterogeneity of beliefs remains hidden within the context of the standard mean-variance analysis. In fact, the mean-variance mutual fund result remains valid within the more general setting with heterogenous beliefs. We also show that distortions of beliefs can be identified separately from preferences for skewness and kurtosis. More specifically, beyond the mean-variance mutual fund one also has a "beliefs-representing" mutual fund where assets are re-weighted in proportion to their associated fads effects. The portfolio is held in a positive quantity by investors whose expectations on risk premiums are scaled by a beliefs factor higher than average. This result coincides with similar findings in the model uncertainty or robustness literature, see for instance Maenhout (2004) and Uppal and Wang (2003). Its empirical implementation follows Anderson, Ghysels, and Juergens (2005a).

Through our theoretical developments we also show that the empirical pricing kernels involve the squared and the cubic market return, the right aggregates to price respectively skewness and kurtosis. However, we also show that when expansions are taken to the next order required to price kurtosis, we find that the dispersion of investors' preferences for skewness gives a nonzero weight to an additional factor ignored by studies based on a representative agent asset pricing model. The fact that heterogeneity of preferences gives rise to additional pricing factors may be related to the general theory of pricing with heterogeneity (see in particular Constantinides and Duffie (1996) and Heaton and Lucas (1995)). Our empirical analysis reveals that this omitted factor is empirically important.

Our closed form formulas for pricing kernels also allowed us to estimate means, variances and covariances of the distribution among investors of preference and beliefs parameters. We used this information in two ways, (1) it allows us to gauge the magnitude of disagreement to have a pricing kernel move inside Hansen-Jagannathan bounds on stochastic discount factors, and (2) it allowed us to obtain empirical estimates of structural dispersion parameters.

We also assessed the empirical evidence pertaining to the pricing kernels for the mean-variance-skewness and mean-variance-skewness-kurtosis asset pricing models. To estimate all the models, we utilized the returns on 20 industry sorted portfolios. Before estimating the parameters of the various specifications, we start with a first assessment based on Hansen-Jagannathan (H-J) bounds and find that for reasonable risk and heterogeneity parameters we find that the implied pricing kernels are empirically plausible and have risk premiums that are empirically reasonable.

When we use GMM-based methods to estimate the structural parameters, we find that the overall model fit is remarkably good, as the J-statistics yield non-rejections of the overidentifying restrictions. The J-statistics also show that the mean-variance-skewness-kurtosis pricing kernel fit is slightly better than the mean-variance-skewness pricing kernel fit. The distance measure and the p-values for the pricing kernels suggest significant improvements moving from the pricing kernel with skewness to the pricing kernel with skewness and kurtosis preferences.

Our empirical findings suggest that the heterogeneity of beliefs plays an important role in the pricing kernel and solves both the equity premium puzzle of Mehra and Prescott (1985) and the risk-free rate puzzle of Weil (1989). We also find that the economic impact of beliefs portfolios is as important as traditional factors, such as the market portfolio.

References

Anderson, Evan, Lars Hansen, and Thomas Sargent, 2003, A Quartet of Semigroups for Model Specification, Robustness, Prices of Risk, and Model Detection, *Journal of the European Economic Association* 1, 68–123.

——— , 2006, Small noise methods for risk sensitive/robust economies, Paper presented at SAMSI Workshop on Model Uncertainty.

Anderson, Evan W., Eric Ghysels, and Jennifer L. Juergens, 2005a, Do heterogeneous beliefs matter for asset pricing?, *Review of Financial Studies* 18, 875–924.

——— , 2005b, The impact of risk and uncertainty on expected returns, Discussion Paper, UNC.

Bansal, R., D. Hsieh, and S. Viswanathan, 1993, A New Approach to International Arbitrage Pricing, *Journal of Finance* 48, 1719–1747.

Chabi-Yo, F., R. Leisen, and E. Renault, 2006, Implications of asymmetry risk for portfolio analysis and asset pricing, Discussion paper, UNC.

Chapman, D., 1997, Approximating the Asset Pricing Kernel, *Journal of Finance* 52, 1383–1410.

Chen, Zengjing, and Larry Epstein, 2002, Ambiguity, risk and asset returns in continuous time, *Econometrica* 70, 1403–1443.

Constantinides, George M., and Darrell Duffie, 1996, Asset pricing with heterogeneous consumers, *Journal of Political Economy* 104, 219–240.

Diether, Karl B., Christopher J. Malloy, and Anna Scherbina, 2002, Differences of opinion and the cross-section of stock returns, *Journal of Finance* 57, 2113–2141.

Dittmar, R. F., 2002, Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from Cross section of Equity Returns, *Journal of Finance* 57, 368–403.

Fang, H., and T. Lai, 1997, Co-kurtosis and capital asset-pricing, *The Financial Review* 32, 293–307.

Gorman, W.M., 1953, Community preference fields, *Econometrica* 21, 63–80.

Hansen, Lars P., and Ravi Jagannathan, 1997, Assessing specification errors in stochastic discount factor models, *Journal of Finance* 52, 557–590.

Hansen, Lars P., and Thomas J. Sargent, 2001, Acknowledging misspecification in macroeconomic theory, *Review of Economic Dynamics* 4, 519–535.

——— , and Thomas D. Tallarini Jr., 1999, Robust permanent income and pricing, *Review of Economic Studies* 66, 873–907.

Hansen, Lars P., Thomas J. Sargent, Gauhar A. Turmuhambetova, and Noah Williams, 2004, Robustness and model misspecification, manuscript.

Harvey, C., and A. Siddique, 2000, Conditional Skewness in Asset Pricing Tests, *Journal of Finance* 55, 1263–1295.

Heaton, John, and Deborah Lucas, 1995, The importance of investor heterogeneity and financial market imperfections for the behavior of asset prices, *Carnegie-Rochester Series on Public Policy* 42, 1–32.

Ingersoll, J., 1987, *Theory of Financial Decision Making* (Rowman & Littlefield Studies in Financial Economics).

Jondeau, E., and M. Rockinger, 2006, Optimal portfolio allocation under higher moments, *European Financial Management* 12, 12–55.

Kimball, M., 1990, Precautionary Saving in the Small and in the Large, *Econometrica* 58, 53–73.

Kogan, Leonid, and Tan Wang, 2002, A simple theory of asset pricing under model uncertainty, Working Paper.

Kraus, A., and R. Litzenberger, 1976, Skewness Preference and the Valuation of Risk Assets, *Journal of Finance* 31, 1085–1100.

Liu, Jun, Jun Pan, and Tan Wang, 2005, An equilibrium model of rare-event premia and its implication for o ption smirks, *Review of Financial Studies* 18, 131–164.

Maenhout, Pascal J., 2004, Robust portfolio rules and asset prices, *Review of Financial Studies* 17, 951–983.

Mehra, Rajnish, and Edward C. Prescott, 1985, The equity premium: A puzzle, *Journal of Monetary Economics* 15, 145–161.

Moskowitz, T., and M. Grinblatt, 1990, Do industries explain momentum?, *Journal of Finance* 54, 1249–1290.

Qu, S., L. Starks, and H. Yan, 2003, Risk dispersion of analyst forecasts and stock returns, mimeo, University of Texas at Austin.

Samuelson, P. A., 1970, The Fundamental Approximation Theorem of Portfolio Analysis in Terms of Mean, Variance and Higher Moments, *Review of Economic Studies* 37, 537–42.

Shefrin, Hersh, 2001, On kernels and sentiment, Mimeo.

Uppal, Raman, and Tan Wang, 2003, Model misspecification and underdiversification, *Journal of Finance* 58, 2465–2486.

Weil, Phillipe, 1989, The equity premium and the risk-free rate puzzle, *Journal of Monetary Economics* 24, 401–421.

Technical Appendix

A Proof of Theorem 1.1 and 1.2

The terminal wealth of agent s is:

$$W_s(\sigma) = q_s \left[R_f + \sum_{i=1}^{n} \omega_{is}(\sigma) \left[R_i^s(\sigma) - R_f \right] \right].$$

For σ given, agent s optimal portfolio choice $(\omega_{is}(\sigma))_{1 \leq i \leq n}$ is characterized by the first-order conditions:

$$E\left[u_s'(W_s(\sigma)) \left[R_i^s(\sigma) - R_f \right] \right] = 0 \quad \text{for} \quad i = 1, ..., n$$

Using the stochastic model (10) for asset returns, we can rewrite these conditions as:

$$\sigma \varphi_{is}(\sigma) = 0$$

where

$$\varphi_{is}(\sigma) = E\left[u_s'(W_s(\sigma)) k_{is}(\sigma) \right]$$

and

$$k_{is}(\sigma) = \sigma a_i(\sigma) + \sigma^2 h_{is}(\sigma) + Y_i.$$

Note that $k_{is}(\sigma)$ is a random variable with zero mean and thus $\varphi_{is}(\sigma) = 0$. If we denote $\varphi_{is}'(\sigma)$ the derivative of $\varphi_{is}(\sigma)$, then according to the results of Samuelson (1970), as extended by Chabi-Yo, Leisen, and Renault (2006), we know that $\omega_{is}(0)$, $\omega_{is}'(0)$ and $\omega_{is}''(0)$ can be characterized by solving respectively $\varphi_{is}'(\sigma) = 0$, $\varphi_{is}''(\sigma) = 0$ and $\varphi_{is}'''(\sigma) = 0$. This gives rise respectively to mean-variance, mean-variance-skewness and mean-variance-skewness-kurtosis pricing. Then we have:

$$\varphi_{is}'(\sigma) = E\left[u_s''(W_s(\sigma)) \frac{dW_s(\sigma)}{d\sigma} k_{is}(\sigma) \right] + E\left[u_s'(W_s(\sigma)) k_{is}'(\sigma) \right]$$

with:

$$\frac{dW_s(\sigma)}{d\sigma} = q_s \sum_{i=1}^{n} \omega_{is}'(\sigma) \left[R_i^s(\sigma) - R_f \right] + q_s \sum_{i=1}^{n} \omega_{is}(\sigma) \frac{dR_i^s}{d\sigma}(\sigma)$$

where

$$\frac{dR_i^s}{d\sigma}(\sigma) = 2\sigma a_i(\sigma) + \sigma^2 a_i'(\sigma) + 3\sigma^2 h_{is}(\sigma) + \sigma^3 h_{is}'(\sigma) + Y_i$$

and

$$k_{is}'(\sigma) = a_i(\sigma) + \sigma a_i'(\sigma) + 2\sigma h_{is}(\sigma) + \sigma^2 h_{is}'(\sigma).$$

Therefore,

$$\varphi_{is}'(0) = E\left[u_s''(q_s R_f) \frac{dW_s}{d\sigma}(0) k_{is}(0) \right] + E\left[u_s'(q_s R_f) k_{is}'(0) \right]$$

with: $\frac{dW_s}{d\sigma}(0)= q_s \sum_{h=1}^{n} \omega'_{hs}(\sigma)Y_h$, $k_{is}(0) = Y_i$, $k'_{is}(0)= a_i(0)$. Therefore, we have:

$$\varphi'_{is}(0) = 0 \Leftrightarrow u''_s(q_sR_f)q_s \sum_{h=1}^{n} \omega'_{hs}(0)E\left[Y_hY_i\right] + u'_s(q_sR_f)a_i(0) = 0.$$

In matrix notation this yields:

$$q_s\Sigma\omega_s(0) - \tau_s a(0) = 0.$$

□

B Proof of Theorem 2.1

Let find the second derivative of $\varphi_{is}(\sigma)$ and $k_{is}(\sigma)$ with respect to σ.

$$\begin{aligned}\varphi''_{is}(\sigma) &= E\left[u'''_s(W_s(\sigma))(\frac{dW_s(\sigma)}{d\sigma})^2 k_{is}(\sigma)\right] + 2E\left[u''_s(W_s(\sigma))\frac{dW_s(\sigma)}{d\sigma}k'_{is}(\sigma)\right] \\ &\quad + E\left[u''_s(W_s(\sigma))\frac{d^2W_s(\sigma)}{d^2\sigma}k_{is}(\sigma)\right] + E\left[u'_s(W_s(\sigma))k''_{is}(\sigma)\right]\end{aligned} \tag{B.1}$$

with:

$$\frac{d^2W_s(\sigma)}{d^2\sigma} = q_s\sum_{i=1}^{n}\omega''_{is}(\sigma)\left[R^s_i(\sigma) - R_f\right] + 2q_s\sum_{i=1}^{n}\omega'_{is}(\sigma)\frac{dR^s_i}{d\sigma}(\sigma) + q_s\sum_{i=1}^{n}\omega_{is}(\sigma)\frac{d^2R^s_i}{d\sigma^2}(\sigma)$$

where:

$$\frac{d^2R^s_i}{d\sigma^2}(\sigma) = 2a_i(\sigma) + 4\sigma a'_i(\sigma) + \sigma^2 a''_i(\sigma) + 6\sigma^2 h'_{is}(\sigma) + 6\sigma h_{is}(\sigma) + \sigma^2 h''_{is}(\sigma) \tag{B.2}$$

and

$$k''_{is}(\sigma) = 2a'_i(\sigma) + \sigma a''_i(\sigma) + 2h_{is}(\sigma) + 4\sigma h'_{is}(\sigma) + \sigma^2 h''_{is}(\sigma).$$

Therefore,

$$\begin{aligned}\varphi''_{is}(0) &= u'''_s(q_sR_f)E\left[(\frac{dW_s}{d\sigma}(0))^2k_{is}(0)\right] + 2u''_s(q_sR_f)E\left[\frac{dW_s}{d\sigma}(0)k'_{is}(0)\right] \\ &\quad + u''_s(q_sR_f)E\left[\frac{d^2W_s}{d\sigma^2}(0)k_{is}(0)\right] + u'_s(q_sR_f)E\left[k''_{is}(0)\right]\end{aligned}$$

with:

$$\begin{aligned}k_{is}(0) &= Y_i, \\ k'_{is}(0) &= a_i(0), \\ k''_{is}(0) &= 2a'_i(0) + 2h_{is}(0) = 2a'_i(0) + 2\alpha_{is}a_i(0),\end{aligned}$$

and

$$
\begin{aligned}
\frac{dW_s}{d\sigma}(0) &= q_s \sum_{i=1}^{n} \omega_{hs}(0) Y_h, \\
\frac{d^2W_s}{d\sigma^2}(0) &= 2q_s \sum_{i=1}^{n} \omega'_{hs}(0) Y_h + 2q_s \sum_{i=1}^{n} \omega_{hs}(0) a_h(0).
\end{aligned}
\tag{B.3}
$$

Therefore, after dividing by $u''_s(q_s R_f)$, we obtain:

$$
\begin{aligned}
\varphi''_{is}(0) = 0 \;\Leftrightarrow\; 0 = & -2\tau_s a'_i(0) - 2\tau_s \alpha_{is} a_i(0) \\
& +2q_s \textstyle\sum_{h=1}^{n} \omega'_{hs}(0) E\left[Y_h Y_i\right] + 2(\rho_s/\tau_s) q_s^2 E\left[\left[\sum_{i=1}^{n} \omega_{hs}(0) Y_h\right]^2 Y_i\right].
\end{aligned}
$$

It is worth noting that:

$$
E\left[\left[\sum_{h=1}^{n} \omega_{hs}(0) Y_h\right]^2 Y_i\right] = E\left[\left[\omega_s^\top(0) Y\right]^2 Y_i\right] = \omega_s^\top(0) E\left[Y Y^\intercal Y_i\right] \omega_s(0) = \omega_s^\top(0) \Gamma_i \omega_s(0)
$$

Then, by substituting the value $\omega_s(0) = (\tau_s/q_s)\Sigma^{-1} a(0)$ given by Theorem 1.2, we have:

$$
E\left[\left[\sum_{h=1}^{n} \omega_{hs}(0) Y_h\right]^2 Y_i\right] = (\frac{\tau_s}{q_s})^2 a^\intercal(0) \Sigma^{-1} \Gamma_i \Sigma^{-1} a(0) = (\frac{\tau_s}{q_s})^2 \xi^\intercal \Gamma_i \xi = (\frac{\tau_s}{q_s})^2 c_i.
$$

Therefore:

$$
\varphi''_{is}(0) = 0 \quad \Leftrightarrow \quad 0 = -\tau_s a'_i(0) - \tau_s \alpha_{is} a_i(0) + q_s \sum_{h=1}^{n} \omega'_{hs}(0) E Y_h Y_i + \rho_s \tau_s c_i.
$$

This can be written as:

$$
q_s \Sigma \omega'_s(0) = \tau_s \left[\rho_s c + a'(0) + \alpha_{.s} \odot a(0)\right]
$$

□

C Proofs of Theorem 2.3 and 2.3

With Assumption 2.2, the slope $\omega'_s(0)$ becomes such that:

$$
\frac{q_s}{\tau_s} \Sigma \omega'_s(0) = \lambda_{0s} a(0) + \sum_{j=1}^{J} \lambda_{js} \alpha^j \odot a(0) + \rho_s c + a'(0).
$$

Hence, the market clearing condition $\sum_{s=1}^{S} q_s \omega'_s(0) = 0$ is tantamount to:

$$
a'(0) = -\overline{\rho} c - \overline{\lambda}_0 a(0) - \sum_{j=1}^{J} \overline{\lambda}_j \alpha^j \odot a(0)
$$

when substituting this formula in $\omega'_s(0)$ above, we obtain the result in Theorem 2.2.

Now, we want to determine a pricing kernel $m(\sigma)$ such that:

$$\begin{aligned} E\left[m(\sigma)R_f\right] &= 1 \\ E\left[m(\sigma)R_i(\sigma)\right] &= 1 \quad \text{for } i = 1,2,...,n \end{aligned}$$

when the probabilistic model for returns is provided by writing a first-order expansion of the risk premium:

$$R_i(\sigma) = R_f + \sigma^2\left[a_i(0) + \sigma a_i'(0)\right] + \sigma Y_i.$$

Then, we look for a zero-mean random variable $Z(\sigma)$ such that:

$$m(\sigma) = \frac{1}{R_f}\left[1 - \sigma Z(\sigma)\right].$$

We have

$$Cov(Y, Z(\sigma)) = a(0) + \sigma a'(0)$$

Let us write

$$a(0) + \sigma a'(0) = (1 - \sigma\overline{\lambda}_0)\Sigma\frac{\overline{\omega}}{\overline{\tau}} - \overline{\rho}\sigma c - \sigma\sum_{j=1}^{J}\overline{\lambda}_j\alpha^j \odot \Sigma\frac{\overline{\omega}}{\overline{\tau}}.$$

We have:

$$\begin{aligned} (1 - \sigma\overline{\lambda}_0)\Sigma\frac{\overline{\omega}}{\overline{\tau}} &= Cov(Y, \frac{1}{\overline{\tau}}\overline{\omega}^\intercal(1 - \sigma\overline{\lambda}_0)Y) \\ -\overline{\rho}\sigma c &= -\overline{\rho}\sigma Cov(Y, (\xi^\intercal Y)^2) = Cov(Y, -\frac{\overline{\rho}\sigma}{\overline{\tau}^2}(\overline{\omega}^\intercal Y)^2) \\ \alpha^j \odot \Sigma\frac{\overline{\omega}}{\overline{\tau}} &= \Sigma\frac{\overline{\omega}}{\overline{\tau}} \odot \alpha^j = Cov(Y, \left[\alpha^j \odot \Sigma\frac{\overline{\omega}}{\overline{\tau}}\right]^\intercal \Sigma^{-1}Y) \end{aligned}$$

Therefore, we choose:

$$Z(\sigma) = \frac{\overline{\omega}^\intercal}{\overline{\tau}}Y(1 - \sigma\overline{\lambda}_0) - \frac{\overline{\rho}\sigma}{\overline{\tau}^2}\left[(\overline{\omega}^\intercal Y)^2 - E(\overline{\omega}^\intercal Y)^2\right] - \sigma\sum_{j=1}^{J}\overline{\lambda}_j(\alpha^j \odot \Sigma\frac{\overline{\omega}}{\overline{\tau}})^\intercal\Sigma^{-1}Y$$

That is:

$$\begin{aligned} Z(\sigma) &= \frac{(1 - \sigma\overline{\lambda}_0)}{\overline{\tau}}\frac{1}{\sigma}(R_M - ER_M) \\ &\quad -\frac{\overline{\rho}\sigma}{\overline{\tau}^2}\left[\frac{1}{\sigma^2}(R_M - ER_M)^2 - \frac{1}{\sigma^2}E(R_M - ER_M)^2\right] - \sigma\sum_{j=1}^{J}\overline{\lambda}_j(\frac{1}{\sigma\overline{\tau}}(R_{Mb}^j - ER_{Mb}^j)). \end{aligned}$$

Therefore,

$$\begin{aligned} m(\sigma) &= \frac{1}{R_f} - \frac{\sigma}{R_f}Z(\sigma) = \frac{1}{R_f} - \frac{1}{R_f\overline{\tau}}(1 - \sigma\overline{\lambda}_0)(R_M - ER_M) \\ &\quad +\frac{\overline{\rho}}{R_f\overline{\tau}^2}\left[(R_M - ER_M)^2 - E(R_M - ER_M)^2\right] + \frac{\sigma}{R_f\overline{\tau}}\sum_{j=1}^{J}\overline{\lambda}_j(R_{Mb}^j - ER_{Mb}^j). \end{aligned}$$

□

D Proof of Theorem 3.1

From expression (B.1), we derive:

$$\begin{aligned}\varphi_{is}^{'''}(\sigma) &= E\left[u_s^{''''}(W_s(\sigma))(\frac{dW_s(\sigma)}{d\sigma})^3 k_{is}(\sigma)\right] \\ &+3E\left[u_s^{'''}(W_s(\sigma))\frac{dW_s(\sigma)}{d\sigma}\frac{d^2W_s(\sigma)}{d\sigma^2}k_{is}(\sigma)\right] + 3E\left[u_s^{'''}(W_s(\sigma))(\frac{dW_s(\sigma)}{d\sigma})^2 k_{is}'(\sigma)\right] \\ &+3E\left[u_s''(W_s(\sigma))\frac{d^2W_s(\sigma)}{d\sigma^2}k_{is}'(\sigma)\right] + 3E\left[u_s''(W_s(\sigma))\frac{dW_s(\sigma)}{d\sigma}k_{is}''(\sigma)\right] \\ &+3E\left[u_s''(W_s(\sigma))\frac{d^3W_s(\sigma)}{d\sigma^3}k_{is}(\sigma)\right] + E\left[u_s'(W_s(\sigma))k_{is}^{'''}(\sigma)\right].\end{aligned}$$

Moreover, we already showed that:

$$\begin{aligned}k_{is}(0) &= Y_i, \\ k_{is}'(0) &= a_i(0), \\ k_{is}''(0) &= 2a_i'(0) + 2h_{is}(0) = 2a_i'(0) + 2\alpha_{is}a_i(0),\end{aligned}$$

since $h_{is}(\sigma) = \alpha_{is}a_i(\sigma)$ and from the expression of $k_{is}''(\sigma)$:

$$k_{is}^{'''}(\sigma) = 3a_i''(\sigma) + \sigma a_i^{'''}(\sigma) + 6h_{is}'(\sigma) + 4\sigma h_{is}''(\sigma) + 2\sigma h_{is}''(\sigma) + \sigma^2 h_{is}^{'''}(\sigma)$$

and thus: $k_{is}^{'''}(0) = 3a_i''(0) + 6\alpha_{is}a_i'(0)$ since $h_{is}'(0) = \alpha_{is}a_i'(0)$. Therefore, after replacement in $\varphi_{is}^{'''}(\sigma)$, we have:

$$\begin{aligned}\varphi_{is}^{'''}(0) &= u_s^{''''}(q_sR_f)E\left[(\frac{dW_s}{d\sigma}(0))^3Y_i\right] + 3u_s^{'''}(q_sR_f)E\left[\frac{dW_s(\sigma)}{d\sigma}\frac{d^2W_s(\sigma)}{d\sigma^2}Y_i\right] \\ &+3u_s^{'''}(q_sR_f)a_i(0)E\left[(\frac{dW_s(\sigma)}{d\sigma})^2\right] + 3u_s''(q_sR_f)a_i(0)E\left[\frac{d^2W_s(\sigma)}{d\sigma^2}\right] \\ &+3u_s''(q_sR_f)(2a_i'(0) + 2\alpha_{is}a(0))E\left[\frac{dW_s(\sigma)}{d\sigma}\right] + u_s''(q_sR_f)E\left[\frac{d^3W_s(\sigma)}{d\sigma^3}Y_i\right] \\ &+u_s'(q_sR_f)(3a_i''(0) + 6\alpha_{is}a_i'(0))\end{aligned}$$

Recall that:

$$\begin{aligned}\tau_s &= -\frac{u_s'(q_sR_f)}{u_s''(q_sR_f)}, \\ \rho_s &= \frac{\tau_s^2}{2}\frac{u_s^{'''}(q_sR_f)}{u_s'(q_sR_f)} \Rightarrow \frac{\rho_s}{\tau_s} = \frac{1}{2}(-\frac{u_s^{'''}(q_sR_f)}{u_s''(q_sR_f)}), \\ \kappa_s &= -\frac{\tau_s^3}{3}\frac{u_s^{''''}(q_sR_f)}{u_s'(q_sR_f)} \Rightarrow \frac{\kappa_s}{\tau_s^2} = \frac{1}{3}\frac{u_s^{''''}(q_sR_f)}{u_s''(q_sR_f)}.\end{aligned}$$

Therefore, after division by $u_s''(q_sR_f)$, we deduce that:

$$
\begin{aligned}
0 = \varphi_{is}'''(0) \quad \Leftrightarrow \quad & \frac{3\kappa_s}{\tau_s^2} E\left[(\frac{dW_s}{d\sigma}(0))^3 Y_i\right] - 6\frac{\rho_s}{\tau_s} E\left[\frac{dW_s}{d\sigma}(0)\frac{d^2W_s}{d\sigma^2}(0)Y_i\right] \\
& -6\frac{\rho_s}{\tau_s} a_i(0) E\left[(\frac{dW_s}{d\sigma}(0))^2\right] + 3a_i(0)E\left[\frac{d^2W_s}{d\sigma^2}(0)\right] \\
& +6(a_i'(0) + \alpha_{is}a(0))E\left[\frac{dW_s}{d\sigma}(0)\right] + E\left[\frac{d^3W_s}{d\sigma^3}(0)Y_i\right] \\
& -3\tau_s(a_i''(0) + 2\alpha_{is}a_i'(0)) = 0
\end{aligned}
$$

From the expression (B.3) and (B.2) of the second derivative of W_s and R_i^s we have:

$$
\begin{aligned}
\frac{d^3W(\sigma)}{d\sigma^3} &= q_s \sum_{i=1}^n \omega_{is}'''(\sigma)\left[R_i^s(\sigma) - R_f\right] + 3q_s \sum_{i=1}^n \omega_{is}''(\sigma)\frac{dR_i^s(\sigma)}{d\sigma} \\
& +3q_s \sum_{i=1}^n \omega_{is}'(\sigma)\frac{d^2R_i^s(\sigma)}{d\sigma^2} + q_s \sum_{i=1}^n \omega_{is}(\sigma)\frac{d^3R_i^s(\sigma)}{d\sigma^3}
\end{aligned}
$$

and $\frac{d^3R_i^s(\sigma)}{d\sigma^3} = 6a_i'(\sigma) + 6h_{is}(\sigma) + \eta_{is}(\sigma)$ with $\eta_{is}(0) = 0$. Using,

$$
\begin{aligned}
R_i^s(0) - R_f &= 0, \\
\frac{dR_i^s}{d\sigma}(0) &= Y_i, \\
\frac{d^2R_i^s}{d\sigma^2}(0) &= 2a_i(0), \\
\frac{d^3R_i^s}{d\sigma^3}(0) &= 6a_i'(0) + 6\alpha_{is}a_i(0),
\end{aligned}
$$

since $h_{is}(0) = \alpha_{is}a_i(0)$, we conclude:

$$
\begin{aligned}
\frac{d^3W}{d\sigma^3}(0) &= 3q_s \sum_{h=1}^n \omega_{hs}''(0)Y_h + 6q_s \sum_{h=1}^n \omega_{hs}'(0)a_h(0) \\
& +6q_s \sum_{h=1}^n \omega_{hs}(\sigma)\left[a_h'(0) + \alpha_{hs}a_h(0)\right].
\end{aligned}
$$

We conclude that:

$$
\begin{aligned}
\varphi_{is}'''(0) = 0 \quad \Leftrightarrow \quad 0 = & \frac{\kappa_s}{\tau_s^2} q_s^3 E\left[(\sum_{h=1}^n \omega_{hs}(0)Y_h)^3 Y_i\right] - \frac{4\rho_s}{\tau_s} q_s^2 E\left[(\sum_{h=1}^n \omega_{hs}(0)Y_h)(\sum_{h=1}^n \omega_{hs}'(0)Y_h)Y_i\right] \\
& -\frac{4\rho_s}{\tau_s} q_s^2 E\left[(\sum_{h=1}^n \omega_{hs}(0)Y_h)(\sum_{h=1}^n \omega_{hs}(0)a_h(0))Y_i\right] - 2\frac{\rho_s}{\tau_s} q_s^2 a_i(0) E\left[(\sum_{h=1}^n \omega_{hs}(0)Y_h)^2\right] \\
& +2q_s a_i(0)\left[\sum_{h=1}^n \omega_{hs}(0)a_h(0)\right] + q_s E\left[(\sum_{h=1}^n \omega_{hs}''(0)Y_h)Y_i\right] - \tau_s\left[a_i''(0) + 2\alpha_{is}a_i'(0)\right].
\end{aligned}
$$

This can be written as:

$$\varphi_{is}^{'''}(0) = 0 \text{ for } i = 1, ..., n \quad \Leftrightarrow \quad \begin{aligned} 0 = & \frac{\kappa_s \tau_s}{q_s} d - 4\frac{\rho_s}{\tau_s} q_s \left[\omega_s^\intercal(0)\Gamma_i \omega_s'(0)\right]_{1\le i\le n} \\ & -\frac{4\rho_s}{\tau_s} q_s(\omega_s^\intercal(0)a(0))\Sigma\omega_s(0) - \frac{2\rho_s}{\tau_s} q_s(\omega_s^\intercal(0)\Sigma\omega_s(0))a(0) \\ & +2(\omega_s^\intercal(0)a(0))a(0) + \Sigma\omega_s''(0) - \frac{\tau_s}{q_s}\left[a''(0) + 2\alpha_{\bullet s}\odot a'(0)\right]. \end{aligned}$$

However, we already know that:

$$\begin{aligned} \omega_s(0) &= \frac{\tau_s}{q_s}\Sigma^{-1}a(0) = \frac{\tau_s}{q_s}\varsigma \\ \omega_s'(0) &= \frac{\tau_s}{q_s}\Sigma^{-1}\left[\rho_s c + a'(0) + \alpha_{.s}\odot a(0)\right] = \frac{\tau_s}{q_s}\left[\rho_s\varsigma_{sk} + \Sigma^{-1}a'(0) + \Sigma^{-1}\varsigma_b^{(s)}\right] \end{aligned}$$

After replacement, we obtain:

$$\varphi_{is}^{'''}(0) = 0 \quad i = 1, \dots, n \quad \Leftrightarrow \quad \begin{aligned} 0 = & \frac{\kappa_s\tau_s}{q_s} d - 4\rho_s\left[\varsigma^\intercal\Gamma_i\omega_s'(0)\right]_{1\le i\le n} \\ & -6\rho_s\frac{\tau_s}{q_s}\frac{\tau_s}{q_s}(\varsigma^\intercal\Sigma\varsigma)(\Sigma\varsigma) + 2(\frac{\tau_s}{q_s}\varsigma^\intercal\Sigma\varsigma)(\Sigma\varsigma) \\ & +\Sigma\omega_s''(0) - \frac{\tau_s}{q_s}\left[a''(0) + 2\alpha_{\bullet s}\odot a'(0)\right] \end{aligned}$$

We conclude that:

$$\begin{aligned} \omega_s''(0) = & -\frac{\kappa_s\tau_s}{q_s}\Sigma^{-1}d + \frac{\tau_s}{q_s}\Sigma^{-1}a''(0) + 2\frac{\tau_s}{q_s}\Sigma^{-1}\alpha_{\bullet s}\odot a'(0) \\ & +2(\varsigma^\intercal\Sigma\varsigma)\frac{\tau_s}{q_s}\left[3\rho_s - 1\right]\varsigma + 4\rho_s\tau_s\frac{\tau_s}{q_s}\Sigma^{-1}\left[\rho_s\varsigma^\intercal\Gamma_i\varsigma_{sk} + \varsigma^\intercal\Gamma_i(\Sigma^{-1}a'(0)) + \varsigma^\intercal\Gamma_i\varsigma_b^{(s)}\right]_{1\le i\le n} \end{aligned}$$

which is the result of Theorem 3.1. □

E Proof of Theorem 3.2

By virtue of Assumption 3.1 we have $J = 1$ and $\lambda_{s0} = 0$, that is $\alpha_{\bullet s} = \lambda_s\, \alpha$. Therefore, the curvature $\omega_s''(0)$ becomes:

$$\begin{aligned} \omega_s''(0) = & -\frac{\kappa_s\tau_s}{q_s}\Sigma^{-1}d + \frac{\tau_s}{q_s}\Sigma^{-1}a''(0) + 2\frac{\tau_s}{q_s}\lambda_s\Sigma^{-1}\alpha_s\odot a'(0) + 2(\varsigma^\intercal\Sigma\varsigma)\frac{\tau_s}{q_s}\left[3\rho_s - 1\right]\varsigma \\ & +4\rho_s\frac{\tau_s}{q_s}\Sigma^{-1}\left[\rho_s\varsigma^\intercal\Gamma_i\varsigma_{sk} + \varsigma^\intercal\Gamma_i(\Sigma^{-1}a'(0)) + \lambda_s\varsigma^\intercal\Gamma_i\varsigma_b\right]_{1\le i\le n} \end{aligned}$$

Hence, the market clearing condition $\sum_{s=1}^{s} q_s\omega_s''(0) = 0$ is tantamount to:

$$\begin{aligned} 0 = & -\overline{\kappa}\Sigma^{-1}d + \Sigma^{-1}a''(0) + 2\overline{\lambda}\Sigma^{-1}(\alpha\odot a'(0)) \\ & +2(3\overline{\rho} - 1)(\varsigma^\intercal\Sigma\varsigma)\varsigma + 4\Sigma^{-1}\left[\overline{\rho^2}\varsigma^\intercal\Gamma_i\varsigma_{sk} + \overline{\rho}\varsigma^\intercal\Gamma_i(\Sigma^{-1}a'(0)) + \overline{\lambda\rho}\varsigma^\intercal\Gamma_i\varsigma_b\right]_{1\le i\le n} \end{aligned}$$

We therefore deduce:

$$\begin{aligned} a''(0) &= \overline{\kappa}d - 2(3\overline{\rho}-1)(\varsigma^\intercal\Sigma\varsigma)\Sigma\varsigma - 2\overline{\lambda}(\alpha\odot a'(0)) \\ &\quad -4\left[\overline{\rho^2}\varsigma^\intercal\Gamma_i\varsigma_{sk} + \overline{\rho}\varsigma^\intercal\Gamma_i(\Sigma^{-1}a'(0)) + \overline{\lambda\rho}\varsigma^\intercal\Gamma_i\varsigma_b\right]_{1\le i\le n} \end{aligned}$$

We now replace $a'(0)$ in the expression above and deduce:

$$\begin{aligned} a''(0) &= \overline{\kappa}d - 2(3\overline{\rho}-1)(\varsigma^\intercal\Sigma\varsigma)\Sigma\varsigma - 2\overline{\lambda}(\alpha\odot a'(0)) \\ &\quad -4(\overline{\rho^2}-\overline{\rho}^2)c_{sk} - 4(\overline{\rho\lambda}-\overline{\rho}\overline{\lambda})c_b \end{aligned}$$

□

F Proof of Theorem 3.3

To be included

G Proof of Theorem 3.4

To be included

Table 1: Data Summary Statistics

This table presents the monthly mean, standard deviation, skewness and kurtosis of returns on 20 industry-sorted portfolios as in Moskowitz and Grinblatt (1990). Portfolios are equally weighted and formed on the basis of two-digit SIC codes. The data is monthly and spans the period December 1981 until December 2005, or a total of 289 monthly returns.

	Mean Return	Standard Deviation	Skewness	Kurtosis
Mining	0.0098	0.0656	-0.1481	5.155
Food & Beverage	0.0136	0.0397	-1.4664	11.0527
Textile & Apparel	0.0085	0.0559	-0.5282	7.1586
Paper Products	0.0124	0.0512	-0.6537	7.0838
Chemical	0.0163	0.0723	0.3543	7.4573
Petroleum	0.012	0.0523	-0.3361	6.0288
Construction	0.0123	0.0551	-0.5444	7.4727
Primary Metals	0.0116	0.0672	-0.0427	5.9363
Fabricated Metals	0.0136	0.0524	-0.553	7.1284
Machinery	0.0123	0.0714	-0.0322	5.2919
Electrical Equipment	0.015	0.0867	0.2555	5.6432
Transport Equipment	0.0117	0.0562	-0.6242	7.1855
Manufacturing	0.0142	0.0638	-0.058	5.664
Railroads	0.0143	0.0518	-0.7714	7.0987
Other Transportation	0.012	0.0535	-0.8337	7.3698
Utilities	0.0123	0.0332	-0.6676	5.6057
Department Stores	0.0123	0.0639	-0.2068	5.1487
Other Retail	0.0114	0.0564	-0.4524	7.0957
Finance, Real Estate	0.0127	0.0364	-0.7203	8.8371
Other	0.0129	0.0703	-0.0188	6.6526

Table 2: Description of Model Parameters and Return Processes

This table provides a summary of the different parameters and returns used in the Mean-Variance-Skewness and Mean-Variance-Skewness-Kurtosis models.

Symbol	Definition	Description
τ_s	Eq. (4)	Risk tolerance coefficient
$1/\tau_s$	Eq. (4)	Arrow-Pratt absolute measure of risk aversion
ρ_s	Eq. (5)	Skew-tolerance coefficient
κ_s	Eq. (10)	Kurtosis-tolerance coefficient
$\sigma\lambda_s$	Eq. (10)	Coefficient of beliefs distortions
$\overline{\omega}$	Def. 2.1	Market portfolio or portfolio selected by investor with average initial wealth, preferences and beliefs
$\overline{\tau}$		Average Risk tolerance coefficient
$1/\overline{\tau}$		Average absolute risk aversion
$\overline{\rho}$	Def. 2.1	Average skew tolerance
$\overline{\sigma\lambda}$	Def. 2.1	Average beliefs distortion
$\overline{\kappa}$	Def. 3.1	Average kurtosis tolerance
$\overline{\rho^2}$	Def. 30	Dispersion of skewness parameters
$\overline{\lambda\rho}$	Def. 30	Covariance of skewness preference parameters and beliefs
$\mathbf{R}_f$		Gross return on riskless asset
$\mathbf{R}_M$		Market Return
$\mathbf{R}_{Mb}$		Payoff on beliefs portfolio (Market beliefs returns)
$\mathbf{R}_i^s(\sigma)$	Eq. (10)	Investor s beliefs of returns
$\sigma h_{is}(\sigma)$	Eq. (10)	Beliefs distortion with respect to objective risk premium $a_i(\sigma)$
ς	Eq. (7)	$\Sigma^{-1}a(0)$ Mean-variance mutual fund
$\varsigma_b^{(s)}$	Def. 1.1	Beliefs portfolio associated with beliefs-distortion function $h_{is}(\sigma)$
ς_b^j	Eq. 12	Beliefs-mimicking mutual funds
$\alpha_{\bullet s}$	Eq. 13	Belief loadings investor s

Table 3: Parameter Estimates Mean-Variance-Skewness-Beliefs Asset Pricing Model

This table reports results of GMM estimation of the Euler equation $E\mathbf{R}_{t+1}m_{t+1}(\theta) = \mathbf{1}$, using the Mean-Variance-Skewness-Beliefs pricing kernel. The coefficients are estimated using a iterated GMM approach and using the Hansen and Jagannathan (1997) weighting matrix. We utilize the returns on 20 industry sorted-sorted portfolios appearing in Table 1. As proxy for the market return, $\mathbf{R}_M$, we use the CRSP value-weighted index. To proxy the "beliefs market return" we utilize the dispersion returns used in Anderson, Ghysels, and Juergens (2005a), which consists of portfolios of stock with large disagreement among analysts. Two portfolios are considered, as there long-run and short-run forecasts, and therefore disagreements on which portfolios are based. Dispersion portfolios are reported with at least 10 and at least 3 analysts coverage respectively. Hence, the parameters are estimated for four beliefs factor returns. For each beliefs factor return, the first line represents the parameter estimates. The second line represents the p-value of the parameter estimates. The third line contains the objective function value, the p-value for the model fit and the J statistics.

	$1/\mathbf{R}_f$	$1/\overline{\tau}$	$\overline{\rho}$	$\sigma\overline{\lambda}$	$1/\mathbf{R}_f$	$1/\overline{\tau}$	$\overline{\rho}$	$\sigma\overline{\lambda}$
	Dispersion portfolios with 10 or more analysts				Dispersion portfolios with 3 or more analysts			
				Iterated GMM				
High LTD	1.0538	13.9978	0.4416	0.5212	1.0861	16.4609	0.3776	0.5663
Top Decile	0.0000	0.0026	0.2455	0.0000	0.0000	0.0007	0.1918	0.0000
	$\delta^2 = 0.0787$	P-value = 0.1576	$J = 22.7492$		$\delta^2 = 0.0711$	P-value = 0.2471	$J = 20.5487$	
High STD	1.0179	19.1168	0.3694	0.8372	1.0250	30.4703	0.1544	0.8673
Top Decile	0.0000	0.0090	0.2193	0.0000	0.0000	0.0011	0.1583	0.0000
	$\delta^2 = 0.0783$	P-value = 0.1619	$J = 22.6235$		$\delta^2 = 0.0539$	P-value = 0.5541	$J = 15.5757$	
				Hansen-Jagannathan Weighting				
High LTD	0.9954	7.9401	0.6506	0.3733	0.9954	7.9730	0.6948	0.3753
Top Decile	0.0000	0.1203	0.5804	0.0556	0.0000	0.1085	0.5553	0.0577
	$\delta^2 = 0.1086$	P-value = 0.0736	$J = 26.049$		$\delta^2 = 0.1078$	P-value = 0.0824	$J = 25.5821$	
High STD	0.9949	21.3078	0.1688	0.8068	0.9946	31.4243	0.0859	0.8404
Top Decile	0.0000	0.0128	0.3777	0.0000	0.0000	0.0070	0.3444	0.0000
	$\delta^2 = 0.0882$	P-value = 0.1327	$J = 23.531$		$\delta^2 = 0.0768$	P-value = 0.4695	$J = 16.7776$	

Table 4: Parameter Estimates Mean-Variance-Skewness-Beliefs Model Implied Pricing Kernel

This table reports results of GMM tests of the Euler equation $E\mathbf{R}_{t+1}m_{t+1}(\theta) = \mathbf{1}$, using the Mean-Variance-Skewness-Beliefs pricing kernel. The coefficients A_0, A_1, A_2, A_3 are computed using estimated structural parameters reported in Table 3. Details of the GMM estimation and the portfolio configurations also appear in that table. The p-values are computed using the Delta method. since A_0, A_1, A_2, A_3 are nonlinear function of the structural preference parameters.

	A_0	A_1	A_2	A_3	A_0	A_1	A_2	A_3
	Dispersion portfolios with 10 or more analysts				Dispersion portfolios with 3 or more analysts			
	Iterated GMM							
High LTD	1.0538	-14.7506	7.6873	91.1863	1.0861	-17.8775	10.1247	111.129
Top Decile	0.0000	0.0056	0.0115	0.1035	0.0000	0.0025	0.0034	0.0635
High STD	1.0179	-19.4592	16.2914	137.4129	1.0250	-31.2323	27.0891	146.9828
Top Decile	0.0000	0.013	0.014	0.0223	0.0000	0.0020	0.0022	0.0233
	Hansen-Jagannathan Weighting							
High LTD	0.9954	-7.9038	2.9506	40.8269	0.9954	-7.9166	2.9710	43.7452
Top Decile	0.0000	0.0545	0.1837	0.2694	0.0000	0.0434	0.1562	0.2263
High STD	0.9949	-21.2001	17.1038	76.2513	0.9946	-31.2555	26.2660	84.3323
Top Decile	0.0000	0.0022	0.0066	0.0382	0.0000	0.0004	0.0012	0.0230

Table 5: Parameter Estimates Mean-Variance-Skewness-Kurtosis-Beliefs Asset Pricing Model

This table reports results of GMM tests of the Euler equation $E\mathbf{R}_{t+1}m_{t+1}(\theta) = \mathbf{1}$, using the Mean-Variance-Skewness-Kurtosis-Beliefs pricing kernel. The coefficients are estimated using a iterated GMM approach. The parameters are estimated for four beliefs factor return. For each beliefs factor returns, the first line represents the parameter estimates. The second line represents the p-value of the parameter estimates. The third line contain the objective function value, the p-value for the model fit and the J statistics. We utilize the returns on 20 industry sorted-sorted portfolios appearing in Table 1. As proxy for the market return, $\mathbf{R}_M$, we use the CRSP value-weighted index. To proxy the "beliefs market return" we utilize the dispersion returns used in Anderson, Ghysels, and Juergens (2005a), which consists of portfolios of stock with large disagreement among analysts. Two portfolios are considered, as there long-run and short-run forecasts, and therefore disagreements on which portfolios are based.

	$1/\mathbf{R}_f$	$1/\overline{\tau}$	$\overline{\rho}$	$\sigma\overline{\lambda}$	$\overline{\kappa}$	$\overline{\rho^2}$	$\overline{\rho\sigma\lambda}$
		Dispersion portfolios with 10 or more analysts					
		Iterated GMM					
High LTD	1.0461	10.9792	0.3495	0.6022	0.6457	0.5307	0.4653
Top Decile	0.0000	0.2544	0.6090	0.1554	0.7713	0.6739	0.3747
	$\delta^2 = 0.0740$	P-value =0.0924	$J = 21.3770$				
High STD	1.0463	23.1502	0.4045	0.6840	0.1381	0.3079	0.3762
Top Decile	0.0000	0.0141	0.0000	0.0056	0.4646	0.1447	0.0917
	$\delta^2 = 0.0628$	P-value =0.2002	$J = 18.1464$				
		Hansen-Jagannathan Weighting					
High LTD	0.9955	3.7680	0.2296	0.6746	23.3731	14.6629	1.7808
Top Decile	0.0000	0.6839	0.9795	0.5721	0.9005	0.8884	0.8080
	$\delta^2 = 0.1042$	P-value = 0.0673	$J = 22.5893$				
High STD	0.9950	16.9593	0.3252	0.9882	0.3162	0.3122	0.4603
Top Decile	0.0000	0.1303	0.2881	0.0766	0.6493	0.4370	0.0755
	$\delta^2 = 0.0840$	P-value = 0.0937	$J = 21.3222$				

Table continued on next page ...

	$1/\mathbf{R}_f$	$1/\overline{\tau}$	$\overline{\rho}$	$\sigma\overline{\lambda}$	$\overline{\kappa}$	$\overline{\rho^2}$	$\overline{\rho\sigma\lambda}$
Dispersion portfolios with 3 or more analysts							
Iterated GMM							
High LTD	1.0684	11.5263	0.1553	0.7015	0.4996	0.6242	0.4755
Top Decile	0.0000	0.1925	0.8623	0.1289	0.7540	0.6183	0.2279
	$\delta^2 = 0.0625$	P-value = 0.2038	$J = 18.0665$				
High STD	1.0562	25.8097	0.3284	0.9361	0.1071	0.2204	0.4440
Top Decile	0.0000	0.0053	0.0001	0.0034	0.4202	0.0753	0.0096
	$\delta^2 = 0.0417$	P-value = 0.6031	$J = 12.0395$				
Hansen-Jagannathan Weighting							
High LTD	0.9955	3.7637	0.5059	0.6645	22.8813	14.1382	1.8612
Top Decile	0.0000	0.6882	0.9533	0.5920	0.9023	0.8883	0.8105
	$\delta^2 = 0.1038$	P-value = 0.0766	$J = 22.0996$				
High STD	0.9946	20.9953	0.2255	1.2689	0.2094	0.2198	0.4492
Top Decile	0.0000	0.0558	0.4686	0.0475	0.5566	0.3355	0.0304
	$\delta^2 = 0.0696$	P-value = 0.4707	$J = 13.7209$				

Table 6: Parameter Estimates Mean-Variance-Skewness-Kurtosis-Beliefs Model Implied Pricing Kernel

This table reports results of GMM tests of the Euler equation $E\mathbf{R}_{t+1}m_{t+1}(\theta) = \mathbf{1}$, using the Mean-Variance-Skewness-Kurtosis-Beliefs pricing kernel. The coefficients A_0 through A_6 are computed using estimated structural parameters reported in Table 5. Details of the GMM estimation and the portfolio configurations also appear in that table. The p-values are computed using the Delta method. since A_0 through A_6 are nonlinear function of the structural preference parameters.

	A_0	A_1	A_2	A_3	A_4	A_5	A_6
				Dispersion portfolios with 10 or more analysts			
				Iterated GMM			
High LTD	1.0461	-11.3237	6.8500	44.0786	-446.9954	1131.278	45.1098
Top Decile	0.0000	0.1608	0.0465	0.7635	0.5817	0.6939	0.6255
High STD	1.0463	-17.7982	16.8747	226.8425	-896.1014	3745.798	-25.4342
Top Decile	0.0000	0.0741	0.0250	0.1969	0.2486	0.2437	0.8784
				Hansen-Jagannathan Weighting			
High LTD	0.9955	-3.7911	2.5042	3.2447	-622.4283	1556.2866	44.3839
Top Decile	0.0000	0.5810	0.3310	0.9781	0.3616	0.5620	0.5717
High STD	0.9950	-17.1470	17.1238	93.0788	-767.2222	2003.6596	-1.7724
Top Decile	0.0000	0.0412	0.0107	0.4730	0.2970	0.4664	0.9890

Table continued on next page ...

	A_0	A_1	A_2	A_3	A_4	A_5	A_6
Dispersion portfolios with 3 or more analysts							
Iterated GMM							
High LTD	1.0684	-14.4159	8.5167	22.0406	-408.6934	1963.7870	93.2852
Top Decile	0.0000	0.0808	0.0258	0.8882	0.6131	0.5294	0.3655
High STD	1.0562	-27.8861	26.1225	231.0267	-971.9555	4087.034	2.4345
Top Decile	0.0000	0.0186	0.0080	0.2580	0.2604	0.2604	0.9904
Hansen-Jagannathan Weighting							
High LTD	0.9955	-3.6809	2.4562	7.1349	-607.2527	1473.7017	39.7041
Top Decile	0.0000	0.5776	0.3065	0.9534	0.3736	0.5767	0.6246
High STD	0.9950	-27.7807	27.3490	98.8687	-963.6039	3109.5760	32.8957
Top Decile	0.0000	0.0052	0.0013	0.4541	0.1935	0.2697	0.8090

Table 7: Beta for the Mean-Variance-Skewness-Beliefs Model

This table reports betas for the 20 industry sorted-sorted portfolios appearing in Table 1. Betas are computed for the full sample with respect to market portfolio $\mathbf{R}_M$, (denoted β_{i2}), the beliefs market return $\mathbf{R}_{Mb}$ (denoted β_{i3} and market squared $\mathbf{R}_M^2$ (denoted β_{i4}). As proxy for the market return, $\mathbf{R}_M$, we use the CRSP value-weighted index. To proxy the "beliefs market return" we utilize the dispersion returns used in Anderson, Ghysels, and Juergens (2005a), which consists of portfolios of stock with large disagreement among analysts. Two portfolios are considered, as there long-run and short-run forecasts, and therefore disagreements on which portfolios are based. Dispersion portfolios are reported with at least 10 and at least 3 analysts coverage respectively.

	10 or more analysts						3 or more analysts					
	High LTD			High STD			High LTD			High STD		
i	β_{i2}	β_{i3}	β_{i4}	β_{i2}	β_{i3}	β_{i4}	β_{i2}	β_{i3}	β_{i4}	β_{i2}	β_{i3}	β_{i4}
Mining	0.5131	0.0695	-2.5597	0.1207	0.5789	-1.9108	0.3473	0.2024	-2.6276	0.0644	0.6293	-2.0694
Food & Beverage	0.459	0.0187	-1.9988	0.2589	0.2697	-1.7105	0.4205	0.0496	-2.0138	0.1785	0.3567	-1.742
Textile Products	0.448	0.1436	-2.5985	0.2704	0.4329	-2.0196	0.3906	0.1907	-2.5826	0.1562	0.5551	-2.0818
Paper Products	0.4883	0.1608	-1.4122	0.2825	0.4933	-0.7554	0.4598	0.185	-1.3748	0.1861	0.5929	-0.8527
Chemical	0.4115	0.4108	-1.6499	0.5816	0.4209	-0.8098	0.2293	0.56	-1.6141	0.3867	0.6381	-0.8047
Petroleum	0.5583	0.0238	-1.679	0.1898	0.4806	-1.1749	0.5023	0.0687	-1.7019	0.1656	0.4959	-1.3242
Construction	0.5637	0.1594	-1.1798	0.3188	0.5381	-0.4797	0.5029	0.2093	-1.1605	0.2182	0.6416	-0.5893
Primary Metals	0.7366	0.1708	-1.2155	0.3801	0.6901	-0.352	0.688	0.2111	-1.1858	0.2864	0.7814	-0.5202
Fabricated Metals	0.5821	0.0652	-1.9849	0.3539	0.3745	-1.5443	0.5174	0.1175	-1.9987	0.2747	0.4573	-1.6134
Machinery	0.5347	0.408	-1.2338	0.5862	0.5595	-0.2582	0.4634	0.4685	-1.1382	0.4372	0.7192	-0.3375
Electrical Equipment	0.3484	0.6875	-1.1218	0.6816	0.6459	0.2258	0.223	0.7939	-0.9636	0.497	0.8449	0.1440
Transport Equipment	0.5648	0.147	-1.7634	0.3296	0.5076	-1.1065	0.5443	0.1647	-1.7261	0.222	0.62	-1.2000
Manufacturing	0.4086	0.3365	-1.9997	0.4665	0.443	-1.2134	0.3079	0.42	-1.9437	0.3248	0.5973	-1.2576
Railroads	0.5873	0.0278	-1.9646	0.301	0.3876	-1.5492	0.5917	0.0246	-1.953	0.2152	0.4778	-1.6178
Other Transportation	0.6493	0.0948	-1.4377	0.415	0.4269	-0.9146	0.601	0.1342	-1.4329	0.3117	0.5366	-0.9832
Utilities	0.4226	-0.0162	-0.1877	0.1992	0.2447	0.04	0.3805	0.0173	-0.2161	0.1618	0.282	-0.0163
Department Stores	0.5757	0.1255	-1.9671	0.3983	0.4051	-1.4343	0.523	0.1687	-1.9545	0.2926	0.5181	-1.4935
Other Retail	0.5184	0.1801	-2.0023	0.39	0.4293	-1.3896	0.4379	0.246	-1.9869	0.2692	0.5593	-1.4454
Finance, Real Estate	0.4415	0.0214	-1.6801	0.2414	0.2738	-1.3848	0.411	0.046	-1.6897	0.1742	0.3453	-1.4281
Other	0.3854	0.4923	-1.7297	0.5236	0.5835	-0.6439	0.2751	0.5849	-1.6276	0.3622	0.7571	-0.7219
Mean	0.5099	0.1863	-1.6683	0.3644	0.4593	-1.0193	0.4408	0.2432	-1.6446	0.2592	0.5703	-1.0977
Standard Deviation	0.0963	0.1864	0.5392	0.1458	0.1201	0.6323	0.1246	0.2124	0.5545	0.1000	0.1459	0.6291

Table 8: Beta for the Mean-Variance-Skewness-Kurtosis-Beliefs Model

This table reports betas for the 20 industry sorted-sorted portfolios appearing in Table 1. Betas are computed for the full sample with respect to market portfolio $\mathbf{R}_M$, (denoted β_{i2}), the beliefs market return $\mathbf{R}_{Mb}$ (denoted β_{i3} the market squared $\mathbf{R}_M^2$ (denoted β_{i4}), the cubic market return $\mathbf{R}_M^3$ (denoted β_{i5}), the cross-product of the market return and a skewness portfolio (denoted β_{i6}), and finally the product of the market return and beliefs portfolio (denoted β_{i7}). As proxy for the market return, $\mathbf{R}_M$, we use the CRSP value-weighted index. To proxy the "beliefs market return" we utilize the dispersion returns used in Anderson, Ghysels, and Juergens (2005a), which consists of portfolios of stock with large disagreement among analysts. Two portfolios are considered, as there long-run and short-run forecasts, and therefore disagreements on which portfolios are based. Dispersion portfolios are reported with at least 10 and at least 3 analysts coverage respectively.

10 or more analysts

	High LTD						High STD					
	β_{i2}	β_{i3}	β_{i4}	β_{i5}	β_{i6}	β_{i7}	β_{i2}	β_{i3}	β_{i4}	β_{i5}	β_{i6}	β_{i7}
i												
Mining	0.4752	0.064	-2.1154	5.6256	6.5297	0.4147	0.0581	0.5839	-3.2151	5.4208	28.2061	3.3343
Food & Beverage	0.4128	0.006	1.5060	7.3802	20.5085	-1.5832	0.2387	0.2500	1.7455	5.1806	0.2922	-3.0527
Textile & Apparel	0.4393	0.1377	-0.1989	-0.4045	27.4506	-1.0667	0.2928	0.4192	0.9675	-2.5456	6.5848	-3.1263
Paper Products	0.4073	0.1489	-0.3687	13.9072	0.1336	0.4213	0.193	0.4839	-0.052	13.4755	7.2127	1.1178
Chemical	0.3487	0.401	-0.5420	12.161	-8.9155	-0.1749	0.5375	0.3998	2.0846	12.5622	-31.3072	-3.1383
Petroleum	0.5311	0.0152	0.9615	-0.7457	54.7942	-0.2000	0.1309	0.4818	-0.9297	-0.2509	75.8987	3.3375
Construction	0.4889	0.1465	0.7111	12.5583	7.4055	-0.1712	0.2585	0.5224	1.6891	10.6513	0.4531	-1.0703
Primary Metals	0.7385	0.1638	2.3124	-1.989	32.6876	-2.0163	0.3563	0.6886	0.2256	-0.5645	36.5442	1.0717
Fabricated Metals	0.5507	0.056	0.7060	6.3851	5.1069	-1.5798	0.3353	0.3619	0.4522	5.1615	-8.5029	-1.9337
Machinery	0.491	0.4024	-1.0532	7.2244	0.053	0.5416	0.5362	0.5512	0.6067	8.593	-2.1847	-0.0509
Electrical Equipment	0.3534	0.6892	-1.6815	-1.0979	-0.8968	0.3577	0.6971	0.6359	1.9581	1.4447	-22.4089	-2.7369
Transport Equipment	0.5155	0.1341	1.6621	8.2866	16.9798	-1.5683	0.2709	0.4948	0.7289	8.6736	12.5004	-0.3653
Manufacturing	0.3873	0.3333	-1.6615	4.6295	-7.2205	-0.1479	0.454	0.4338	0.0192	5.1878	-19.8365	-1.6653
Railroads	0.5312	0.0172	-0.0835	8.2748	16.983	-0.1874	0.2545	0.377	0.0903	6.2322	15.7827	-0.2764
Other Transportation	0.5978	0.0835	1.0922	7.7799	18.6543	-0.7479	0.384	0.4111	1.7724	5.6559	6.1816	-1.906
Utilities	0.375	-0.0268	2.1965	9.4216	0.185	-1.2115	0.1692	0.2306	2.0893	7.4589	-12.2924	-1.9119
Department Stores	0.6028	0.1212	1.7571	-9.2078	57.9511	-1.8619	0.4597	0.3913	2.5179	-11.1659	31.1225	-3.9435
Other Retail	0.5141	0.1752	0.1854	-1.8306	31.7752	-0.839	0.2181	0.2698	-0.6284	1.7045	18.9248	0.2968
Finance, Real Estate	0.4172	0.016	-0.4493	2.6555	16.8331	-0.1569	0.2181	0.2698	-0.6284	1.7045	18.9248	0.2968
Other	0.3828	0.492	-1.7083	1.5260	-8.3386	-0.2112	0.5328	0.5746	0.7248	2.7721	-26.8514	-2.4158
Mean	0.4780	0.1788	0.1614	4.6270	14.4330	-0.5994	0.3392	0.4490	0.7000	4.1341	6.6274	-1.0340
Standard Deviation	0.0990	0.1892	1.3706	5.8990	19.1674	0.8167	0.1616	0.1231	1.3314	5.8460	25.0972	2.0695

Table continued on next page ...

3 or more analysts

	High LTD						High STD					
	β_{i2}	β_{i3}	β_{i4}	β_{i5}	β_{i6}	β_{i7}	β_{i2}	β_{i3}	β_{i4}	β_{i5}	β_{i6}	β_{i7}
i												
Mining	0.3102	0.1994	-2.0964	5.3297	5.2352	0.2438	-0.0033	0.6384	-3.8104	6.7987	17.3205	3.3615
Food & Beverage	0.3693	0.0409	1.4196	7.4662	17.7856	-1.572	0.1674	0.3272	1.6299	5.9768	-5.8936	-3.1059
Textile & Apparel	0.375	0.1857	-0.2798	0.7694	23.301	-1.0022	0.1796	0.5357	0.4294	-0.972	-0.4457	-2.7273
Paper Products	0.3697	0.1779	-0.0247	14.7104	-1.8776	0.1309	0.0971	0.5831	-0.5243	14.6187	-2.7115	1.1467
Chemical	0.1414	0.5549	-1.1944	15.6439	-16.5596	0.4446	0.3294	0.6196	0.899	14.595	-33.2031	-1.7038
Petroleum	0.4775	0.0613	1.24	-1.1604	53.8953	-0.5309	0.1181	0.4887	-1.0384	0.442	63.6286	2.7253
Construction	0.419	0.2013	0.9222	13.3782	4.3838	-0.3803	0.1588	0.6222	1.1546	12.2292	-9.4072	-0.8948
Primary Metals	0.6843	0.2044	2.5107	-0.7283	26.6717	-2.2141	0.2677	0.7765	-0.1597	0.9321	20.6211	0.693
Fabricated Metals	0.481	0.1113	0.6591	6.769	1.4418	-1.6096	0.2589	0.441	0.1785	6.2466	-16.2631	-1.9251
Machinery	0.4027	0.4647	-0.758	10.2165	-6.0784	0.385	0.38	0.7141	-0.2544	10.317	-10.8793	0.5444
Electrical Equipment	0.2008	0.794	-1.587	4.5279	-12.186	0.4817	0.4984	0.8401	0.7839	3.9518	-30.3645	-1.6446
Transport Equipment	0.4846	0.1556	1.7515	9.4593	12.2699	-1.6414	0.1686	0.6025	0.3604	9.8641	1.3781	-0.4621
Manufacturing	0.2727	0.4179	-1.6058	7.1622	-13.3928	-0.1512	0.3047	0.5915	-0.7214	6.7336	-26.2675	-1.0389
Railroads	0.5317	0.0176	0.0953	8.2631	16.3875	-0.3553	0.1743	0.4622	-0.1774	7.1119	7.1157	-0.3736
Other Transportation	0.5422	0.1266	1.0741	8.3828	15.8872	-0.7604	0.2862	0.5141	1.4312	6.8476	-2.632	-1.8664
Utilities	0.3288	0.011	2.0306	9.234	-1.4481	-1.138	0.138	0.2625	2.0855	8.076	-18.4826	-2.1477
Department Stores	0.5473	0.1629	1.755	-8.1537	52.8846	-1.8865	0.3633	0.4919	2.215	-9.7997	22.8958	-3.8505
Other Retail	0.4271	0.2413	0.1622	-0.503	27.2384	-0.8331	0.2862	0.5431	0.609	-1.5242	8.9443	-1.8769
Finance, Real Estate	0.3847	0.0422	-0.4804	2.6915	16.1688	-0.1582	0.1517	0.3393	-0.9598	2.4248	13.8806	0.4313
Other	0.2524	0.5842	-1.6754	5.5161	-17.2156	-0.1216	0.3575	0.754	-0.3353	5.046	-33.9638	-1.4298
Mean	0.4001	0.2377	0.1959	5.9487	10.2396	-0.6334	0.2341	0.5574	0.2898	5.4958	-1.7365	-0.8073
Standard Deviation	0.1300	0.2143	1.3850	5.8879	20.3130	0.8283	0.1192	0.1507	1.3429	5.8591	23.2753	1.8509

Table 9: Market Price of Risk for the Mean-Variance-Skewness-Beliefs Model

3 and More Analyst	Iterated GMM			HJ Weighting		
	$\Lambda_2 \times 10^{-2}$	$\Lambda_3 \times 10^{-2}$	$\Lambda_4 \times 10^{-2}$	$\Lambda_2 \times 10^{-2}$	$\Lambda_3 \times 10^{-2}$	$\Lambda_4 \times 10^{-2}$
High LTD	1.9503	0.2595	-0.2368	1.3518	0.9663	-0.1163
Top Decile	0.7866	0.9314	0.1075	0.3703	0.5179	0.0571
High STD	2.5393	0.3084	-0.2734	2.3520	0.0824	-0.1761
Top Decile	0.9068	0.7948	0.1156	0.4988	0.4294	0.0580

10 and More Analyst	Iterated GMM			HJ Weighting		
	$\Lambda_2 \times 10^{-2}$	$\Lambda_3 \times 10^{-2}$	$\Lambda_4 \times 10^{-2}$	$\Lambda_2 \times 10^{-2}$	$\Lambda_3 \times 10^{-2}$	$\Lambda_4 \times 10^{-2}$
High LTD	1.8417	0.4326	-0.2095	1.3324	0.8865	-1.1125
Top Decile	0.7382	0.8499	0.1031	0.3683	0.5548	0.0578
High STD	2.0760	0.4711	-0.2487	1.9297	0.1324	-0.1511
Top Decile	0.8488	0.7538	0.1076	0.4423	0.4626	0.0567

Table 10: Market Price of Risk for the Mean-Variance-Skewness-Kurtosis-Beliefs Model

Iterated GMM						HJ Weighting					
$\Lambda_2 \times 10^{-2}$	$\Lambda_3 \times 10^{-2}$	$\Lambda_4 \times 10^{-2}$	$\Lambda_5 \times 10^{-2}$	$\Lambda_6 \times 10^{-2}$	$\Lambda_7 \times 10^{-2}$	$\Lambda_2 \times 10^{-2}$	$\Lambda_3 \times 10^{-2}$	$\Lambda_4 \times 10^{-2}$	$\Lambda_5 \times 10^{-2}$	$\Lambda_6 \times 10^{-2}$	$\Lambda_7 \times 10^{-2}$
3 or more analysts											
High LTD Top Decile											
1.6982	0.2908	-0.1900	0.0327	0.0087	-0.2661	1.2347	0.9646	-0.1246	0.0278	0.0068	-0.1708
0.7802	1.0563	0.1156	0.0224	0.0076	0.1442	0.4258	0.5252	0.0580	0.0108	0.0039	0.0831
High STD Top Decile											
2.7087	0.5485	-0.2842	0.0478	0.0117	-0.2555	1.2347	0.9646	-0.1246	0.0278	0.0068	-0.1708
1.0467	0.9146	0.1361	0.0277	0.0087	0.1505	0.5383	0.4445	0.0601	0.0105	0.0034	0.0615
10 or more analysts											
High LTD Top Decile											
1.7657	0.5791	-0.2069	0.0373	0.0106	-0.2525	1.2261	0.9345	-0.1219	0.0270	0.0065	-0.1636
0.8184	1.0310	0.1288	0.0255	0.0084	0.1479	0.4250	0.5682	0.0585	0.0111	0.0041	0.0805
High STD Top Decile											
2.1529	0.5230	-0.2334	0.0390	0.0096	-0.2049	1.8622	0.1179	-0.1609	0.0343	0.0083	-0.1560
0.9057	0.8333	0.1211	0.0241	0.0078	0.1346	0.4839	0.4755	0.0583	0.0101	0.0034	0.0603