BANK OF CANADA
BANQUE DU CANADA

Working Paper/Document de travail
2008-16

On Portfolio Separation Theorems with Heterogeneous Beliefs and Attitudes towards Risk

by Fousseni Chabi-Yo, Eric Ghysels, and Eric Renault

www.bank-banque-canada.ca

Bank of Canada Working Paper 2008-16

May 2008

On Portfolio Separation Theorems with Heterogeneous Beliefs and Attitudes towards Risk

by

Fousseni Chabi-Yo[1], Eric Ghysels[2], and Eric Renault[3]

[1]Financial Markets Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
fchabiyo@bankofcanada.ca

[2]Author to whom correspondence should be addressed:
Department of Economics
University of North Carolina at Chapel Hill
CB# 3305, Gardner Hall
Chapel Hill, NC 27599-3305
eghysels@unc.edu
and
Department of Finance
Kenan-Flagler Business School

[3]Department of Economics
University of North Carolina at Chapel Hill
and
CIRANO, CIREQ, Montréal
renault@email.unc.edu

ISSN 1701-9397

Acknowledgements

We thank Eugene Fama, Dietmar Leisen, James Heckman as well as seminar participants at the Bank of Canada, CREST, the University of Chicago, UBC and the 2006 EC^2 Conference in Rotterdam for helpful comments.

Abstract

The early work of Tobin (1958) showed that portfolio allocation decisions can be reduced to a two stage process: first decide the relative allocation of assets across the risky assets, and second decide how to divide total wealth between the risky assets and the safe asset. This so called two-fund separation relies on special assumptions on either returns or preferences. Tobin (1958) analyzed portfolio demand in a mean-variance setting. We revisit the fund separation in settings that allow not only for heterogeneity of preferences for higher order moments, but also for heterogeneity of beliefs among agents. To handle the various sources of heterogeneity, beliefs, and preferences, we follow the framework of Samuelson (1970) and its recent generalization by 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. Besides the standard Sharpe-Lintner CAPM mutual fund separation we obtain additional mutual funds called beliefs portfolios, pertaining to heterogeneity of beliefs, a skewness portfolio similar to Kraus and Litzenberger (1976), beliefs about skewness portfolios with design quite similar to beliefs portfolios, a kurtosis portfolio, and finally portfolio heterogeneity of the preferences for skewness across investors in the economy as well as its covariation with heterogeneity of beliefs. These last two mutual funds are called cross-co-skewness portfolio and cross-co-skewness-beliefs portfolios. Under various circumstances related to return distribution characteristics, cross-agent heterogeneity and market incompleteness, some of these portfolios disappear.

JEL classification: C52, D58, G11, G12
Bank classification: Financial markets; Market structure and pricing

Résumé

Dans ses premiers travaux, Tobin (1958) a montré que les décisions d'allocation d'un portefeuille peuvent se résumer en deux étapes : l'investisseur décide d'abord de la répartition des actifs au sein de la composante risquée de son portefeuille, puis de la proportion de la richesse totale à investir respectivement dans les actifs risqués et dans l'actif sans risque. Ce « théorème de séparation à deux fonds » repose sur l'adoption d'hypothèses concernant soit les rendements, soit les préférences. Tobin (1958) analysait la demande d'actifs à l'aide d'un modèle moyenne-variance. Les auteurs réexaminent ici la question dans un cadre qui autorise non seulement l'hétérogénéité des préférences pour les moments d'ordre supérieur à un, mais aussi celle des croyances des investisseurs. Pour prendre en compte les diverses croyances, préférences et autres sources d'hétérogénéité existantes, ils empruntent l'approche de Samuelson (1970) ainsi que la généralisation qu'ils en ont proposée récemment dans Chabi-Yo, Leisen et Renault (2006). Leur approche générale permet d'établir, pour un niveau de risque donné, les parts optimales de la richesse qu'un agent optimisateur ayant des préférences à l'égard de l'espérance, de la variance, de l'asymétrie et de l'aplatissement de la distribution des rendements investirait dans chaque titre. Outre le portefeuille de marché classique construit à partir du modèle d'équilibre des actifs financiers de Sharpe et Lintner, les auteurs arrivent ainsi à générer : 1) différents portefeuilles qui sont fonction des croyances des investisseurs; 2) un portefeuille lié à l'asymétrie de la distribution des rendements et similaire à celui de Kraus et Litzenberger (1976); 3) des portefeuilles tributaires des croyances des agents concernant cette asymétrie et analogues aux premiers portefeuilles décrits ci-dessus; 4) un portefeuille tenant à l'aplatissement excessif de la distribution des rendements; 5) un autre associé à l'hétérogénéité des préférences à l'égard de l'asymétrie; et 6) des portefeuilles dont l'existence s'explique par la covariation de cette hétérogénéité avec celle des croyances. Dans diverses circonstances déterminées par les caractéristiques de la distribution des rendements, l'hétérogénéité des agents et l'incomplétude des marchés, certains de ces portefeuilles disparaissent.

Classification JEL : C52, D58, G11, G12
Classification de la Banque : Marchés financiers; Structure de marché et fixation des prix

Introduction

The early work of Tobin (1958) showed that portfolio allocation decisions can be reduced to a two stage process: first decide the relative allocation across risky assets, and second decide how to divide total wealth between risky assets and the riskless asset. This so called two-fund separation relies on special assumptions on either returns or tastes. Tobin (1958) analyzed portfolio demand in a mean-variance setting. Two-fund separation has indeed been examined in great detail in the CAPM model, see for example Black (1972). The Tobin approach assumed a quadratic utility function. Since then there have been two approaches to portfolio separation. Cass and Stiglitz (1970) provide conditions on agents' preferences that ensure two-fund separation whereas Ross (1978) presents conditions on asset return distributions under which two-fund separation holds. Finally, Russell (1980) presents a unified approach of Cass and Stiglitz (1970) and Ross (1978).

In this paper we revisit fund separation theorems and the conditions regarding preferences and return distributions under which they hold. First, we go beyond the usual representative agent utility models by allowing not only for heterogeneity of preferences for higher order moments, but also for heterogeneity of beliefs among agents. To handle the various sources of heterogeneity, beliefs and preferences that exist, 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[1], 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 and the mutual funds that relate to them. We use an analogy with the model uncertainty literature to model heterogeneous beliefs. Namely, we assume that J factors of potential beliefs distortions linearly enter the risk premiums as perceived by investors. Each investor s has personal beliefs differing from the reference model measured via J factor loadings that are specific to agent s. While the factor loadings cancel out - by definition - for the reference model, we do not assume that the average across investors sums up to zero.

To describe the main findings of our paper, consider the case of n primitive assets that have mutually uncorrelated returns. Our analysis yields various mutual funds, which we can describe as follows.

First, as in a standard Sharpe-Lintner CAPM, any investor s holds a share of the market portfolio. Equilibrium prices are such that each risky asset enters the composition of the market portfolio proportional to its risk premium per unit of variance. In equilibrium, the size of the share held by investor s is proportional to his/her coefficient of risk tolerance.

Second, the J factors of beliefs distortions introduce J additional mutual funds called *beliefs portfolios.* These portfolios appear as a trimming of the market portfolio, in the simplest case where a beliefs factor can be represented by an indicator function that selects a specific subset of assets. In this simple case, the beliefs portfolio only gives a non-zero weight to the assets in the specific subset still determined by the risk premiums per unit of

[1]By infinitely small risk, we refer to the perturbation approximation approach proposed by Samuelson (1970)

variance. To carry this analysis further, one can think of J factors of beliefs distortions, each selecting mutually exclusive subsets of assets. While the J beliefs portfolios are in zero net supply, investor s holds beliefs portfolio $j, = 1, \ldots, J$, proportional to his/her risk tolerance times the spread between his/her specific factor loadings and the cross-sectional average of loadings across all investors.

Third, the presence of skewness in the asset return distribution yields an additional mutual fund, called the *skewness portfolio,* similar to Kraus and Litzenberger (1976). The skewness portfolio attributes weights that are proportional to the co-skewness (per unit of variance) of the n assets with the market portfolio. These weights are zero when the multivariate return distribution is symmetric - as in the Gaussian case. Otherwise, even though the skewness portfolio is in zero net supply, the weights of the portfolio held by individual investor s are in proportion to his/her risk tolerance times the deviation from the cross-sectional average of his/her specific skewness tolerance.

Fourth, beyond the skewness portfolio, investor s may also hold J *beliefs about skewness* portfolios. The design of such portfolios is quite similar to that of beliefs portfolios, namely they tilt the skewness portfolio on the basis of the J beliefs factors, like the beliefs portfolio trims the market portfolio. Hence, investor s holds $j, = 1, \ldots, J$, beliefs about skewness portfolios with weights proportional to his/her risk tolerance times the spread between his/her specific factor loadings and the cross-sectional average of loadings across all investors.

Fifth, excess kurtosis in the return distribution that is non-uniform across assets introduces a *kurtosis portfolio.* The weights of this portfolio are determined by the co-kurtosis of the n assets with the market portfolio. This portfolio coincides with the market portfolio - and hence de facto disappears - when the excess kurtosis is identical across assets, say zero as in the Gaussian case. Under all other circumstances, the kurtosis portfolio, although in zero net aggregate supply, is held by individual investor s with weights that are in proportion to his/her risk tolerance times the deviation from the cross-sectional average of his/her individual kurtosis tolerance.

Sixth, heterogeneity of preferences for skewness across investors in the economy as well as heterogeneity of beliefs may also yield additional mutual funds. At a general level we know that heterogeneity of preferences and beliefs matters when there is some form of market incompleteness. In our case, this incompleteness exists because the squared market return cannot be perfectly hedged with linear portfolios and because investors have belief distributions which are not uniform across assets. Under these circumstances additional mutual funds emerge and they relate to the heterogeneity of preferences and beliefs. These mutual funds are called *cross-co-skewness* portfolio and *cross-co-skewness-beliefs* portfolios. Under perfect hedging of squared market returns the cross-co-skewness portfolio coincides with the kurtosis portfolio. The J cross-co-skewness-beliefs portfolios coincide with the skewness portfolio when beliefs distortions are uniform across assets. The cross-co-skewness portfolio is again in zero net supply, and investor s holds such a portfolio with weights proportional to the risk tolerance times the spread between his/her specific contribution to the cross-sectional variance of skewness tolerance and the population variance of skewness tolerance. Likewise, the J cross-co-skewness-beliefs portfolios are also in zero net supply and similar to the cross-co-skewness portfolio, except that the cross-co-skewness is computed with respect to the j^{th} beliefs portfolio, $j = 1, \ldots, J$. The asset allocation of investor s in such portfolios is similar to the cross-co-skewness portfolio choice, except that cross-sectional covariance between skew tolerance and the j^{th} factor loading of beliefs distortions is used instead of the variance of skewness tolerance.

It is the purpose of the paper to rigorously derive the individual asset demands in equilibrium as linear combinations of the aforementioned mutual funds. Our derivations do not require that asset returns are uncorrelated in the cross-section. Compared to the existing literature we make several contributions. First, we show that preferences for higher moments result in various additional mutual funds beyond the skewness and kurtosis portfolios introduced by several authors (Kraus and Litzenberger (1976), Ingersoll (1987) and Dittmar (2002)), due to the market incompleteness and the heterogeneity of skewness tolerance across investors. Our results are reminiscent of Constantinides and Duffie (1996) who present a model where market incompleteness also yields an additional pricing factor related to a measure of dispersion - i.e. the cross-sectional variance - of heterogeneity across agents. Similar to our analysis, Constantinides and Duffie (1996) introduce an additional mutual fund that is has zero aggregate supply. Second, similar to Maenhout (2004) we find that pessimistic (optimistic) belief distortions may be observationally equivalent to an upward (downward) bias in risk aversion. However, we do show that there is room for separate identification of beliefs distortions and risk aversion - similar to what Uppal and Wang (2003) find. Namely, beliefs distortions that are not uniform across assets, provides a vehicle to identify effective risk aversion coefficients that differ across subsets of assets. Third, we also contribute to the literature on investor heterogeneity and portfolio choice. Namely, we point out that there are mutual fund portfolios generated by the interaction between heterogeneous preferences and beliefs, as they result in a covariance (across the population of investors) between skew tolerance and loadings to belief distortion factors.

The paper is organized as follows. In section 1 we start with a brief description of **?**'s small noise expansion framework and its recent generalization by 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 Section 2 we introduce a Mean-Variance-Skewness asset pricing model and derive equilibrium portfolio allocations with heterogeneous beliefs and preferences. Section 3 takes the analysis a step further as we study a Mean-Variance-Skewness-Kurtosis environment. Section 4 concludes the paper.

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.[2] Chabi-Yo, Leisen,

[2]Consider, for instance, an asset with mean μ and variance σ^2. By small risk, we refer to the case where the santard deviation of this asset is small, that is $\sigma \mapsto 0$. When σ is small, the second moment will count

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 small noise expansion of the first order optimality conditions one step beyond the Newton 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) and introduce heterogeneous preferences into Samuelson's small noise expansion setting.[3] 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 $(\omega_{is})_{1 \le i \le n} \in \mathbb{R}^n$ that maximize her/his expected utility for a given initial wealth invested q_s :

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

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

where, $\mathbf{R}_f$ is the gross return on the riskless asset and the solution is denoted by $(\omega_{is}(\sigma))_{1 \le i \le n}$ 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 $\mathbf{R} = (\mathbf{R}_i)_{1 \le i \le n}$ defines the objective joint probability distribution of interest, which is specified by the following decomposition:

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

for less and less. However, when σ is small but not limitingly small, the third moment of skewness will still count along with the mean and variance of this asset

[3]See also Judd and Guu (2001) and Anderson, Hansen, and Sargent (2006) for recent work on small noise expansions

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 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. In this paper, we are interested in 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_iY_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 non-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.

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, with 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) analysis 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 behaviour 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 $u_s^{''''}(q_s\mathbf{R}_f)$.

Since small noise expansions are based on the local behaviour of the utility function around zero risk, we follow Judd and Guu (2001) and characterize preferences by their derivatives at the terminal wealth in a hypothetical risk-free environment evaluated at the value $q_s\mathbf{R}_f$. In particular, preferences are 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.[4] 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 nonnegative, following the literature on preferences for higher order moments (Chapman (1997), Dittmar (2002), Harvey and Siddique (2000), Jondeau and Rockinger (2006) and Guidolin and Timmermann (2006)). Note that the positivity of skew tolerance is also supported by the literature on prudence (see in particular Kimball (1990)).

Next, we show that the standard mean-variance formulas always provide the local first-order 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), (5) and (6). 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 neighbourhood 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 risky portfolio (the same across all s), dubbed the mean-variance mutual fund and defined by the coefficients of the vector:

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

[4]Note that preferences may be labeled heterogeneous (τ_s, ρ_s and κ_s differ across agents) whenever agents differ by their utility function or by their initial wealth q_s.

The individual risk tolerance coefficient τ_s defines the respective weights of the two mutual funds in the portfolio of investor s.

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, 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 payoff 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. In other words, (1) an assumption of joint symmetry of the probability distribution of returns, which is in particular fulfilled in the 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, and (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 the usual 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.[5] 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.[6]

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 be in the neighborhood of zero when the economy is in the neighborhood of zero risk ($\sigma = 0$). Hence, the beliefs of investor s, for $s = 1, ...S$, are defined by the following stochastic model for returns:

$$\mathbf{R}_i^s(\sigma) = \mathbf{R}_f + (1 + \sigma\alpha_{is})\sigma^2 a_i(\sigma) + \sigma Y_i \tag{10}$$

where $\sigma\alpha_{is}$ represents relative beliefs distortions with respect to the *"objective"* risk premium $\sigma^2 a_i(\sigma)$. Note that the notion of objective risk premiums will only be well defined once we implement the model empirically, namely when the pricing kernel will be identified and defined with respect to the so called *"objective"* probability measure (see Chabi-Yo, Ghysels, and Renault (2007) for further details). 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). Conversely, this degree of freedom is not key for our results. There is a purpose for all the beliefs-based mutual funds we introduce, irrespective of the average value of beliefs distortions (which may or may not be zero).

We do impose some constraints, however, on the distortions of beliefs of the average investor. In particular, note that in equation (10) the beliefs distortion function $\sigma\alpha_{is}$ is proportional to σ, and this ensures that the price of risk $a_i(0)$ is not modified by beliefs distortions in the neighborhood of zero risk.[7] Consequently, heterogeneity of beliefs will not be identified through asset demands and equilibrium prices in the limit case as σ goes to zero.

[5]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.

[6]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.

[7]A more general formulation would consist of replacing α_{is} by a function of σ, say $h_{is}(\sigma)$. Such a formulation would only result in higher order effects that do not appear empirically relevant. Therefore, to simplify notation, we do not pursue the more general setting.

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 appearing in Theorem 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, which 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 $(a'_i(0) + \alpha_{is}a_i(0))$, 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 explored 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 be related to beliefs portfolios, with the purpose of hedging the heterogeneity of beliefs. In particular, we can define:

Definition 1.1 *A beliefs-distortion function α_{is} 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 $\alpha_{\bullet s} \equiv (\alpha_{is})_{(1 \le i \le n)}$.

Note that a beliefs portfolio $\varsigma_b^{(s)}$ does not coincide (up to a scaling factor) with the mean-variance portfolio $\varsigma = \Sigma^{-1} a(0)$, unless the components of $\alpha_{\bullet s}$ are all equal. Hence, we need to go beyond the mean-variance portfolio, whenever for some agent s, the (α_{is}) are not equal. This implies that we require additional *"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 $(\alpha_{\bullet s})_{(1 \le s \le S)}$ 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 $\alpha^j \equiv (\alpha_i^j)_{(1 \le i \le n)}$, also independent of $\mathbf{1}$. It will be shown in the next section that this beliefs heterogeneity will yield J beliefs-mimicking mutual funds:

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

The above spanning condition implies that we have a system of beliefs loadings, $(\lambda_{sj})_{(1 \le j \le J)}$, for each investor s such that:

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

The decomposition (13) produces a model for returns reminiscent of a factor structure in risk premia:

$$R_i^s(\sigma) = R_f + \sigma^2 a_i(\sigma)[1 + \sigma(\lambda_{s0} + \sum_{j=1}^{J} \lambda_{sj}\alpha_i^j)] + \sigma Y_i \tag{14}$$

The idea that investor's expectations regarding payoffs for different assets $i = 1, \ldots, n$, can be linearly decomposed into payments α_i^j, $j = 1, \ldots, J$. The loadings λ_{sj}, $j = 1, \ldots, J$ represent the unit price of these characteristics according to the expectations of investor s. A similar approach is used in labour and product market models. For example, the decomposition is similar to the Gorman-Lancaster model of earnings analyzed by Heckman and Scheinkman (1987), where J separate productive attributes of an individual i on the labour market have prices at different dates. This setup has its underpinnings in consumer demand theory. Gorman (1959) and Lancaster (1966) introduced models where various unobservable characteristics of goods yield utility with different weights (or prices).[8]

The J asset characteristics will potentially result in J mutual funds sufficient to characterize equilibrium portfolios for all investors and equation (12) represents the most general

[8]It should be noted parenthetically that Gorman (1959) used additive separability for the purpose of aggregating consumer demands. In the present context an additive separable setup is more natural and convenient.

definition of such J beliefs-mimicking portfolios. It will be useful to specialize the discussion to the particular case where characteristics $j, = 1, \ldots, J$ are binary attributes. Each asset i is endowed with a subset $J(i)$ of these attributes:

$$\alpha_i^j = \begin{cases} 1 & j \in J(i) \\ & \\ 0 & Otherwise \end{cases} \tag{15}$$

Asset i is not exposed to characteristics $j \notin J(i)$. As a consequence, investor's expectations about asset i are not affected by such characteristics j, very much like expected future earnings for an individual i are not affected by an attribute j not in the endowment of i in a labour market factor model.[9] Then, the beliefs portfolio ς_b^j in (12) appears as a trimming of the mean-variance portfolio of Theorem 1.2 since it only assigns a non-zero weight to the assets i such that $j \in J(i)$.

2 Mean-Variance-Skewness-Beliefs Mutual Funds

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.

2.1 The individual investor problem

A small noise expansion one order beyond the Newton 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:*

$$q_s[\omega_s(0) + \sigma\omega_s^{'}(0)] = \tau_s\Sigma^{-1}[(\mathbf{1} + \sigma\alpha_{\bullet s})] \odot a(0) + \sigma\tau_s\Sigma^{-1}[\rho_s c + a^{'}(0)]. \tag{16}$$

Proof: See Appendix B.

Recall that identification issues may potentially arise, since one may expect that distortions of beliefs may be confounded with preferences for skewness in the slopes of optimal portfolio shares, i.e. $\omega_{is}^{\prime}(0)$. Theorem 2.1 clearly disentangles the respective roles of preferences for skewness and distortion of beliefs. Besides the mean-variance portfolio $\varsigma = \Sigma^{-1}a(0)$,

[9] Cunha, Heckman, and Navarro (2005) use such an expectations model to separate uncertainty from heterogeneity in a life cycle model of earnings. In Chabi-Yo, Ghysels, and Renault (2007) this factor model allows us to identify separately beliefs and preferences.

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 $\varsigma_{[3]} = \Sigma^{-1}c$. 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 the return on $\varsigma_{[3]}$. Hence, while derivative assets with nonlinear payoffs may be in practice a way to trade skewness, the skewness portfolio $\varsigma_{[3]}$ appears in our framework as the best way to replicate the nonlinear payoff $(\varsigma'\mathbf{R})^2$ when 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:*

$$\alpha_{\bullet s} = \lambda_{s0}\mathbf{1} + \sum_{j=1}^{J} \lambda_{sj}\alpha^j \qquad \forall s = 1, \ldots, S. \tag{17}$$

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:

$$\begin{aligned} ER^s(\sigma) - R_f\mathbf{1} &= \sigma^2[(1+\sigma\lambda_{s0})a(0) + \sigma a'(0)] \\ &+\sigma^2[\sigma\sum_{j=1}^{J}\lambda_{sj}\alpha^j \odot a(0)], \end{aligned} \tag{18}$$

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 (18). 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 imply 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:*

$$\overline{\tau} = \frac{1}{S}\sum_{s=1}^{S}\tau_s, \quad \overline{\rho} = \frac{\sum_{s=1}^{S}\tau_s\rho_s}{\sum_{s=1}^{S}\tau_s}, \quad \overline{\lambda}_j = \frac{\sum_{s=1}^{S}\tau_s\lambda_{sj}}{\sum_{s=1}^{S}\tau_s} \qquad j = 0,1,\ldots,J \tag{19}$$

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:

$$a(0) \;=\; \Sigma\overline{\omega}/\overline{\tau} \qquad a'(0) \;=\; -\overline{\rho}c - \overline{\lambda}_0(\alpha \odot a(0)) - \textstyle\sum_{j=1}^{J}\overline{\lambda}_j\alpha^j \odot a(0) \tag{20}$$

To summarize, we have the following theoretical result:

Theorem 2.2 *Let Assumptions 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:*

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

Proof:The proof is obtained by replacing (20) in Theorems 2.1.
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 of the 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 investor's risk tolerance.

- 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 $\varsigma_{[3]} = \Sigma^{-1}c$. The corresponding portfolio is held in a positive quantity by investor 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 portfolios defined by shares proportional to $\varsigma_b^j = \Sigma^{-1}[\alpha^j \odot a(0)]$. It is held in a positive quantity by investor 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 assets i are heterogeneous with respect to the attribute j.

Note that, by definition, the skewness portfolio $\varsigma_{[3]}$ 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:

$$(\tau_s/\overline{\tau})[\mathbf{1} + \sigma(\lambda_{0s} - \overline{\lambda}_0)]\overline{\omega} \tag{21}$$

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 asset 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) (his formula (17) on 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.

To conclude, as far as preference 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 combined with asymmetries in asset payoffs. 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. It will be shown in the next subsection that higher order features of beliefs distortions are helpful to uncover that investors may not realize that their beliefs are biased. Brunnermeier and Parker (2005) provide an alternative justification for the persistence of biased beliefs. Forward-looking agents have higher current *"felicity"* if they are optimistic. Interestingly enough, Brunnermeier and Parker (2005) also find that overoptimistic investors both overestimate their returns ($\lambda_{sj} > \overline{\lambda}_j$) and have a strong preference for skewness ($\rho_s > \overline{\rho}$). The two effects are clearly disentangled by our Theorem 2.2. Nevertheless, it is true that optimism and preference for skewness produce similar additional asset demands with respect to the market portfolio. However, since they give rise to different mutual funds, it leaves room for separate identification of these two effects.

3 Mean-Variance-Skewness-Kurtosis-Beliefs Mutual Funds

Having introduced the Mean-Variance-Skewness portfolio separation theorem, 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. First, in Subsection 3.1 we analyze the individual investor problem, before deriving in Subsection 3.2 the implications for equilibrium allocations and prices.

3.1 Individual Asset Demand

A small noise expansion two steps beyond the quadratic approximation yields a price for skewness and kurtosis, or a mean-variance-skewness-kurtosis framework using the simplified model (14) 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 small noise expansion up to the fourth degree:

Theorem 3.1 *Assume the setting of Theorem 1.2, with preferences to 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:*

$$\begin{array}{rcl} \omega_s(0) & = & (\tau_s/q_s)\varsigma \\ \omega_s'(0) & = & (\tau_s/q_s)[\rho_s\varsigma_{[3]} + \varsigma_b^s + \Sigma^{-1}a'(0)] \\ \omega_s''(0)/2 & = & (\tau_s/q_s)[\varsigma_b^{s\prime} + \Sigma^{-1}a''(0)/2] \\ & & -\kappa_s(\tau_s/q_s)\varsigma_{[4]}/2 + (\tau_s/q_s)(3\rho_s - 1)(\varsigma^\intercal\Sigma\varsigma)\varsigma \\ & & +2(\tau_s/q_s)\Sigma^{-1}\rho_s\left[\rho_s(\varsigma^\intercal\Gamma_i\varsigma_{[3]}) + \varsigma^\intercal\Gamma_i\varsigma_b^s + \varsigma^\intercal\Gamma_i\varsigma'\right]_{1\le i\le n} \end{array} \tag{22}$$

where:

$$\begin{array}{rclrcl} \varsigma & = & \Sigma^{-1}a(0) & \varsigma_{[3]} & = & \Sigma^{-1}c \\ \varsigma_{[4]} & = & \Sigma^{-1}d & \varsigma_b^s & = & \Sigma^{-1}[\alpha_{\bullet s}\odot a(0)] \\ \varsigma' & = & \Sigma^{-1}a'(0) & \varsigma_b^{s\prime} & = & \Sigma^{-1}[\alpha_{\bullet s}\odot a'(0)] \end{array} \tag{23}$$

Proof: See Appendix C.

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 $\varsigma = \Sigma^{-1}a(0)$ 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 demand of investor s do not involve skewness, kurtosis or preferences for them.

Third, note that the skewness portfolio $\varsigma_{[3]} = \Sigma^{-1}c$ is now augmented with a kurtosis portfolio $\varsigma_{[4]} = \Sigma^{-1}d$. Its coefficients in the asset demand of investor s do not involve beliefs distortions and are simply proportional to the intensities ρ_s and κ_s respectively of the s skew and kurtosis tolerances. The coefficients c_i and d_i measure the contributions of asset i in the skewness and kurtosis respectively 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 the preference of investor s for positive skewness ρ_s, and 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 of the mean-variance mutual fund, $(\varsigma^\intercal \mathbf{R})^2$, on the vector of risky asset returns, $\mathbf{R} = (\mathbf{R}_i)_{1 \leq i \leq n}$, is an affine function of the skewness portfolio $\varsigma_{[3]}$. This yields an alternative interpretation of the skewness portfolio. Not surprisingly, following similar arguments, one can show that the affine regression of the cubic return $(\varsigma^\intercal \mathbf{R})^3$ on $\mathbf{R}$ is an affine function of the kurtosis portfolio $\varsigma_{[4]}$. Hence, the skewness and kurtosis portfolios are respectively the best mimicking portfolios for the nonlinear payoffs $(\varsigma^\intercal \mathbf{R})^2$ and $(\varsigma^\intercal \mathbf{R})^3$.

Again, in the case of joint normality of returns, similar to the result in section 2, the skewness portfolio $\varsigma_{[3]}$ vanishes while the kurtosis portfolio $\varsigma_{[4]}$ is simply proportional to the mean-variance portfolio ς.

Beyond the skewness portfolio $\varsigma_{[3]} = \Sigma^{-1} c$ 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 :

$$c_{ib}^s = \varsigma^\intercal \Gamma_i \varsigma_b^s = Cov[(\varsigma^\intercal Y)((\varsigma_b^s)^\intercal Y), Y_i] \tag{24}$$

Along similar lines, we also define a cross-co-skewness measure between the mean-variance mutual fund ς and the skewness portfolio $\varsigma_{[3]}$ as well as a cross-co-skewness measure with the "differentiated" portfolio $\varsigma' = \Sigma^{-1} a(0)$:

$$c_{i[3]} = \varsigma^\intercal \Gamma_i \varsigma_{[3]} = Cov[(\varsigma^\intercal Y)(\varsigma_{[3]}^\intercal Y), Y_i]$$

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

It is worth recalling briefly that all these cross-co-skewnesses are zero - like the co-skewnesses c_i - whenever the primitive asset return distributions are symmetric ($\Gamma_i = 0$). Moreover, Theorem 3.1 shows that cross-co-skewnesses give rise to additional portfolios, $\Sigma^{-1} c_b^s$, $\Sigma^{-1} c_{[3]}$ 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 to define some aggregate quantities in order to characterize the equilibrium outcome. 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:*

$$\overline{\kappa} = \sum_{s=1}^{S} \kappa_s \frac{\tau_s}{\sum \tau_s} \tag{25}$$

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* $s = 1, \ldots, S : \alpha_{is} = \lambda_s \ \alpha_i$.

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:

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

with $c_{[3]} = (c_{i[3]})_{1 \le i \le n}$ with $c_{i[3]} = \varsigma^{\intercal}\Gamma_i\varsigma_{[3]}$ and $c_b = (c_{ib})_{1 \le i \le n}$ with $c_{ib} = \varsigma^{\intercal}\Gamma_i\varsigma_b$ and average quantities defined as:

$$\overline{\rho^2} = \sum_{s=1}^{S} \rho_s^2 \frac{\tau_s}{\sum_{s=1}^{S} \tau_s} \qquad \overline{\lambda\rho} = \sum_{s=1}^{S} \rho_s \lambda_s \frac{\tau_s}{\sum_{s=1}^{S} \tau_s} \tag{26}$$

Therefore, to summarize, we have established the following:

Theorem 3.2 *The vector of asset risk premiums in equilibrium as given in (2),* $\sigma^2 a(\sigma)$, *admits a second order Taylor expansion in the neighborhood of zero risk characterized by:*

$$\begin{aligned} a(0) \ &= \ \Sigma\overline{\omega}/\overline{\tau} \\ a^{'}(0) \ &= \ -\overline{\lambda}(\alpha \odot a(0)) - \overline{\rho}c \\ a''(0) \ &= \ \overline{\kappa}d - 2(3\overline{\rho} - 1)(\overline{\omega}^{\intercal}\Sigma\overline{\omega})\Sigma\overline{\omega}/\overline{\tau}^3 \\ &\quad -4(\overline{\rho^2} - \overline{\rho}^2)c_{[3]} - 4(\overline{\rho\lambda} - \overline{\lambda}\overline{\rho})c_b \\ &\quad +2\overline{\lambda}\overline{\rho}(\alpha \odot c) + 2\overline{\lambda}^2[\alpha \odot (\alpha \odot a(0))] \end{aligned} \tag{27}$$

Proof: See Appendix D.

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 into 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 $\varsigma_{[3]}$ is to track the squared market return $(\varsigma^\intercal\mathbf{R})^2$, namely the linear regression of $(\varsigma^\intercal\mathbf{R})^2$ on $\mathbf{R}$ is an affine function of $\varsigma_{[3]}$. 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 $\varsigma_{[3]}$. However, markets are not complete with regards to this *"quadratic risk"*. There is a non-zero residual risk, as in general:

$$Var[(\varsigma^\intercal\mathbf{R})^2] \geq Var[(\varsigma_{[3]}^\intercal\mathbf{R})]$$

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 only yields a compensation for co-kurtosis coefficients d_i, while we find in addition to this a compensation for cross-skewness coefficients $c_{i[3]} = Cov[(\varsigma^\intercal Y)(\varsigma_{[3]}^\intercal Y), Y_i]$. These coefficients measure the contribution of asset i to the aggregate risk:

$$\sum_{i=1}^{n} \varsigma_i c_{i[3]} = Cov[(\varsigma^\intercal Y)^2, (\varsigma_{[3]}^\intercal Y)] = E[(\varsigma^\intercal Y)^2(\varsigma_{[3]}^\intercal Y)]$$

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 $(\varsigma_{[3]}^\intercal\mathbf{R})$.

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.

When both aggregate beliefs distortion $\overline{\lambda}$ and aggregate skew tolerance $\overline{\rho}$ are non-zero, two additional priced factors emerge, as shown in Theorem 3.2. These factors are characterized by $\alpha \odot c$ and $\alpha \odot \alpha \odot a(0)$. 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:

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

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

$$\sum_{i=1}^{n} \varsigma_i c_{ib} = Cov[(\varsigma^{\intercal}Y)^2, (\varsigma_b^{\intercal}Y)] = Cov[(\varsigma_{[3]}^{\intercal}Y), (\varsigma_b^{\intercal}Y)]$$

since the hedging error $(\varsigma^{\intercal}Y)^2$ - $(\varsigma_{[3]}^{\intercal}Y)$ 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 premia 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 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:*

$$q_s\omega_s(\sigma) = q_s[\omega_s(0) + \sigma\omega_s^{'}(0) + \sigma^2\omega_s^{''}(0)/2] \tag{28}$$

$$\begin{aligned}
q_s\omega_s(0) &= \tau_s\varsigma \\
q_s\omega_s^{'}(0) &= \tau_s(\rho_s - \overline{\rho})\varsigma_{[3]} + \tau_s\left(\lambda_s - \overline{\lambda}\right)\varsigma_b \\
\frac{1}{2}q_s\omega_s^{''}(0) &= 3\tau_s(\rho_s - \overline{\rho})(\varsigma^{\intercal}\Sigma\varsigma)\varsigma - \frac{1}{2}\tau_s(\kappa_s - \overline{\kappa})\varsigma_{[4]} \\
&\quad -\tau_s\overline{\rho}\left(\lambda_s - \overline{\lambda}\right)\varsigma_{b[3]} - \tau_s\overline{\lambda}\left(\lambda_s - \overline{\lambda}\right)\varsigma_{[b,b]} \\
&\quad +2\tau_s\left[\rho_s(\rho_s - \overline{\rho}) - \left(\overline{\rho^2} - \overline{\rho}^2\right)\right]\varsigma_{[3,3]} \\
&\quad +2\tau_s\left[\rho_s\left(\lambda_s - \overline{\lambda}\right) - \left(\overline{\rho\lambda} - \overline{\rho}\overline{\lambda}\right)\right]\varsigma_{[3,b]}
\end{aligned}$$

where: $\varsigma_{[3]} = \Sigma^{-1}c$, $\varsigma_{[3,3]} = \Sigma^{-1}\left[\varsigma^{\intercal}\Gamma_i\varsigma_{[3]}\right]_{1\le i\le n}$, $\varsigma_b = \Sigma^{-1}[\alpha \odot (\Sigma\varsigma)]$, $\varsigma_{[b,b]} = \Sigma^{-1}[\alpha \odot (\alpha \odot \Sigma\varsigma)]$, $\varsigma_{b[3]} = \Sigma^{-1}[\alpha \odot c]$, $\varsigma_{[3,b]} = \Sigma^{-1}[\varsigma^{\intercal}\Gamma_i\varsigma_b]_{1\le i\le n}$, *and* $\varsigma_{[4]} = \Sigma^{-1}d$.

Proof: See Appendix E.

A particular investor s will hold in equilibrium shares of several mutual funds, where the shares depend on the spread between his/her preference/beliefs characteristics and the economy-wide averages summarized in Table 1. Note that Theorem 3.3 is a mutual funds separation theorem which displays eight mutual funds in equilibrium. Since, for simplicity, we considered a one-factor beliefs structure, we have introduced five additional mutual funds with respect to the mean-variance-skewness-beliefs pricing of section 2 (see Theorem 2.2). To summarize, equilibrium individual asset demands $\omega_s(\sigma)$, represented by their second order expansion appearing in equation (28), involve the following mutual funds in addition to the mean-variance portfolio $\varsigma = \overline{\omega}/\overline{\tau}$

- The beliefs portfolio $\varsigma_b = \Sigma^{-1}[\alpha \odot \Sigma\varsigma]$
- The skewness portfolio $\varsigma_{[3]} = \Sigma^{-1}c$ with $c = (c_i)_{1\leq i\leq n}$ where $c_i = \varsigma^\intercal\Gamma_i\varsigma = Cov\left((\varsigma^\intercal Y)^2, Y_i\right)$
- The kurtosis portfolio $\varsigma_{[4]} = \Sigma^{-1}d$ with $d = (d_i)_{1\leq i\leq n}$ where $d_i = Cov\left((\varsigma^\intercal Y)^3, Y_i\right)$
- The beliefs-about-skewness portfolio $\varsigma_{b[3]} = \Sigma^{-1}[\alpha \odot c]$
- The beliefs-about-beliefs portfolio $\varsigma_{[b,b]} = \Sigma^{-1}[\alpha \odot (\alpha \odot \Sigma\varsigma)]$
- The cross-coskewness portfolio $\varsigma_{[3,3]} = \Sigma^{-1}c_{[3]}$ with $c_{[3]} = \left(c_{i[3]}\right)_{1\leq i\leq n}$ where $c_{i[3]} = \varsigma^\intercal\Gamma_i\varsigma_{[3]} = Cov\left((\varsigma^\intercal Y)\left(\varsigma_{[3]}^\intercal Y\right), Y_i\right)$
- The cross co-skewness beliefs portfolio $\varsigma_{[3,b]} = \Sigma^{-1}c_b$ with $c_b = (c_{ib})_{1\leq i\leq n}$ where $c_{ib} = \varsigma^\intercal\Gamma_i\varsigma_b = Cov\left((\varsigma^\intercal Y)(\varsigma_b^\intercal Y), Y_i\right)$

Note that the beliefs-about-beliefs portfolio $\varsigma_{[b,b]}$ coincides with the beliefs portfolio ς_b when the latter appears as a trimming of the market portfolio through an asset-attributes structure like (15).

Where do these additional funds come from and how do they relate to the existing ones appearing in the mean-variance-skewness-beliefs pricing model? We devote the remainder of this section to this topic.

Beliefs and skewness/kurtosis portfolios

Similar to the case covered by Theorem 2.2, we note that the mean-variance mutual fund ς is augmented with the beliefs portfolio ς_b and the skewness portfolio $\varsigma_{[3]}$. The skewness portfolio $\varsigma_{[3]}$ becomes immaterial when the joint return distribution is symmetric ($\Gamma_i = 0$) while the beliefs portfolio ς_b differs from the mean-variance portfolio ς only when the beliefs distortions α_i are nonuniform across the n assets. The beliefs portfolio (respectively the skewness portfolio) is held in a positive quantity by investor s whose beliefs loading λ_s (respectively skewness tolerance ρ_s) is higher than the average.

Similarly, the kurtosis portfolio $\varsigma_{[4]} = \Sigma^{-1}d$ collapses in $3\left(Var\left(\varsigma^\intercal Y\right)\right)\varsigma$ in the case of joint normal returns (see (9)). In the case of returns with heterogeneous degrees of leptokurticity, $\varsigma_{[3]}$ is an additional mutual fund (no longer proportional to the mean-variance portfolio ς), held in a positive quantity by investor s whose kurtosis tolerance κ_s is smaller than the average.

The effect of heterogeneity

The higher-order expansion may give rise to four additional mutual funds: two of them are associated with non-zero population averages of skewness tolerance and beliefs dispersion, respectively, while the others are associated with non-zero population variances of these characteristics of preferences and beliefs.

On both theoretical and empirical grounds, non-zero population variances are the most important effects. First, the fact that skewness tolerances ρ_s are heterogeneous gives rise to the cross-coskewness portfolio $\varsigma_{[3,3]}$. It is held in positive quantity by investor s whose spread-to-mean skewness tolerance has a positive impact on the cross-sectional variance $\overline{\rho^2} - \overline{\rho}^2$ of skew tolerances. Typically, this effect of heterogeneity of preferences for skewness is ignored by mean-variance-skewness-kurtosis pricing models based on a representative investor (see e.g. Dittmar (2002)). Heterogeneity of preferences for skewness has an additional effect when it implies a non-zero cross-sectional covariance with beliefs distortion λ_s. The cross-co-skewness-beliefs portfolio $\varsigma_{[3,b]}$ is held in positive quantity by investor s whose skew-tolerance ρ_s is not only positive but also has a positive impact on the population covariance between skew-tolerance and beliefs dispersion.

Finally, when the population mean $\overline{\rho}$ of skew tolerance (respectively the population mean $\overline{\lambda}$ of beliefs distortion) is non-zero, an investor s whose beliefs distortion λ_s is smaller than the average will hold a positive quantity of the beliefs-about-skewness portfolio $\varsigma_{b[3]}$ (respectively the beliefs-about-beliefs portfolio $\varsigma_{[b,b]}$). While higher order beliefs patterns are captured by $\varsigma_{b[3]}$ and ς_{bb}, the cross-coskewness beliefs portfolio $\varsigma_{[3,b]}$ and even more so the cross-coskewness portfolio $\varsigma_{[3,3]}$ may play a more interesting role. Their role may be understood through market incompleteness, similar to how Constantinides and Duffie (1996) find a role for the cross-sectional variance of individual consumption growth when individual consumption risk cannot be hedged. Indeed, similar to their results, we find a role for the cross sectional variance $\overline{\rho^2} - \overline{\rho}^2$ of skewness tolerances and to a lesser extent for cross-sectional covariance $\overline{\rho\sigma\lambda} - \overline{\rho}\,\overline{\sigma\lambda}$, because some specific risks cannot be hedged. More precisely, the cross-coskewness portfolio $\varsigma_{[3,3]}$ has a specific role because for some i:

$$Cov\left(\left(\varsigma^{\intercal}Y\right)\left(\varsigma_{[3]}^{\intercal}Y\right), Y_i\right) \neq Cov\left(\left(\varsigma^{\intercal}Y\right)^3, Y_i\right)$$

because the regression $\left(\varsigma_{[3]}^{\intercal}Y\right)$ of $\left(\varsigma^{\intercal}Y\right)^2$ on Y does not coincide with $\left(\varsigma^{\intercal}Y\right)^2$. As already mentioned in section 2.1, the skewness portfolio $\varsigma_{[3]}^{\intercal}Y$ is the best, albeit imperfect, hedge of squared market return with primitive asset returns. It is precisely because this hedge is imperfect that the cross-co-skewness portfolio and the kurtosis portfolio are two different mutual funds. Similarly, this is because the two hedging portfolios of $\left(\varsigma^{\intercal}Y\right)^2$ and $\left(\varsigma^{\intercal}Y\right)\left(\varsigma_b^{\intercal}Y\right)$ by primitive returns are different than the cross-coskewness-beliefs and the skewness portfolio.

4 Conclusions

In this paper we revisited fund separation theorems and the conditions under which preferences and return distributions hold. We followed 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.

Using a small noise expansion two orders beyond the quadratic approximation, our analysis yields various components to equilibrium asset demand beyond those in the standard CAPM. The portfolios represent return characteristics and attitudes towards risk pertaining to skewness and kurtosis as well as beliefs characteristics among investors in the economy.

In related work, Chabi-Yo, Ghysels, and Renault (2007), we study the asset pricing implications of our framework, both from a theoretical and empirical perspective. Our theoretical analysis reveals that the empirical pricing kernels involve the squared and the cubic market return as well as the dispersion of investors' preferences for skewness. 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. 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 whose weights depend on the dispersion across investors' preferences and beliefs, and the interactions between them, which are analogous to the mutual fund representation characterized in the present paper. From an empirical perspective, such a structural pricing kernel allows us to statistically identify characteristics of the population heterogeneity like $\overline{\rho^2}$ - $\overline{\rho}^2$ and $\overline{\rho\lambda}$ - $\overline{\rho}\overline{\lambda}$ while using only aggregate returns data.

References

Anderson, E., E. Ghysels, and J. Juergens (2005a). Do heterogeneous beliefs matter for asset pricing? *Review of Financial Studies 18*, 875–924.

Anderson, E., E. Ghysels, and J. Juergens (2005b). The impact of risk and uncertainty on expected returns. Discussion Paper, UNC.

Anderson, E., L. Hansen, and T. 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.

Anderson, E., L. Hansen, and T. Sargent (2006). Small noise methods for risk sensitive/robust economies. Paper presented at SAMSI Workshop on Model Uncertainty.

Black, F. (1972). Capital Market Equilibrium with Restricted Borrowing. *Journal of Business 45*, 444–455.

Brunnermeier, M. and J. Parker (2005). Optimal expectations. *American Economic Review*, 1092–1118.

Cass, D. and J. Stiglitz (1970). The Structure of Investor Preferences and Asset Returns, and Separability in Portfolio Allocation: A Contribution to the Pure Theory of Mutual Funds. *Journal of Economic Theory 2*, 122–160.

Chabi-Yo, F., E. Ghysels, and E. Renault (2007). Disentangling the effects of heterogeneous beliefs and preferences on asset prices. Discussion Paper, Bank of Canada and UNC.

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, Z. and L. Epstein (2002). Ambiguity, risk and asset returns in continuous time. *Econometrica 70*, 1403–1443.

Constantinides, G. M. and D. Duffie (1996, April). Asset pricing with heterogeneous consumers. *Journal of Political Economy 104*(2), 219–240.

Cunha, F., J. Heckman, and S. Navarro (2005). Separating heterogeneity from uncertainty in life cycle earnings. *Oxford Economic Papers 57*, 191–261.

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. (1959). Separable utility and aggregation. *Econometrica 27*, 469–481.

Guidolin, M. and A. Timmermann (2006). International Asset Allocation under Regime Switching, Skew and Kurtosis Preferences . Working paper, UCSD.

Hansen, L. P. and T. J. Sargent (2001). Acknowledging misspecification in macroeconomic theory. *Review of Economic Dynamics 4*(3), 519–535.

Hansen, L. P., T. J. Sargent, and T. D. Tallarini, Jr. (1999). Robust permanent income and pricing. *Review of Economic Studies 66*(4), 873–907.

Hansen, L. P., T. J. Sargent, G. A. Turmuhambetova, and N. 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, J. and D. 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.

Heckman, J. and J. Scheinkman (1987). The importance of bundling in a gorman-lancaster of earnings. *Review of Economic Studies 54*, 243–255.

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.

Judd, K. and S. Guu (2001). Asymptotic Methods for Asset Market Equilibrium Analysis. *Economic Theory 18*, 127–157.

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

Kogan, L. and T. 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.

Lancaster, K. J. (1966). A new approach to consumption theory. *Journal of Political Economy 74*, 132–157.

Liu, J., J. Pan, and T. Wang (2005). An equilibrium model of rare-event premia and its implication for o ption smirks. *Review of Financial Studies 18*(1), 131–164.

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

Ross, S. A. (1978). Mutual Fund Separation in Financial Theory - The Separating Distributions. *Journal of Economic Theory 17*, 254–286.

Russell, T. (1980). Portfolio Separation: The Analytic Case. *Economics Letters 6*, 59–66.

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.

Tobin, J. (1958). Liquidity Preferences as Behavior Towards Risk. *Review of Economic Studies 25*, 65–86.

Uppal, R. and T. Wang (2003). Model misspecification and underdiversification. *Journal of Finance 58*, 2465–2486.

Technical Appendix

A Proof of Theorems 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 $(\omega_{is}(\sigma))_{1 \le i \le n}$ 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

$$k_{is}(\sigma) = \sigma a_i(\sigma)[1 + \sigma \alpha_{is}].$$

Note that $k_{is}(0)$ is a random variable with zero mean and thus $\varphi_{is}(0) = 0$. Let us denote $\varphi_{is}'(\sigma)$, $\varphi_{is}''(\sigma)$, $\varphi_{is}'''(\sigma)$ as the consecutive derivatives 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 $\varphi_{is}'(0) = 0$, $\varphi_{is}''(0) = 0$ and $\varphi_{is}'''(0) = 0$. 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

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

and

$$k_{is}'(\sigma) = (1 + \sigma \alpha_{is})[a_i(\sigma) + \sigma a_i'(\sigma)] + \sigma \alpha_{is} a_i(\sigma).$$

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

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

In matrix notation this yields:

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

B Proof of Theorem 2.1

The second derivative of $\varphi_{is}(\sigma)$ and $k_{is}(\sigma)$ with respect to σ is:

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

with:

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

where:

$$\frac{d^2 R_i^s}{d\sigma^2}(\sigma) = 2a_i(\sigma) + 4\sigma a'_i(\sigma) + \sigma^2 a''_i(\sigma) + 6\sigma^2 \alpha_{is} a'_i(\sigma) + 6\sigma\alpha_{is} a_i(\sigma) + \sigma^3 \alpha_{is} a''_i(\sigma) \tag{B.2}$$

and

$$k''_{is}(\sigma) = 2\alpha_{is}[a_i(\sigma) + \sigma a'_i(\sigma)] + (1 + \sigma\alpha_{is})(2a'_{is}(\sigma) + \sigma\alpha''_{is}(\sigma)).$$

Therefore,

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

with:

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

and

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

Therefore, after dividing by $u_s''(q_sR_f)$, we obtain:

$$\begin{array}{rcl} \varphi_{is}''(0) = 0 \;\Leftrightarrow\; 0 & = & -2\tau_s a_i'(0) - 2\tau_s\alpha_{is}a_i(0) \\ & & +2q_s\sum_{h=1}^n \omega_{hs}'(0)E\left[Y_hY_i\right] - 2(\rho_s/\tau_s)q_s^2 E\left[\left[\sum_{h=1}^n \omega_{hs}(0)Y_h\right]^2 Y_i\right]. \end{array}$$

It is worth noting that:

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

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:

$$\varphi_{is}''(0) = 0 \quad\Leftrightarrow\quad 0 = -\tau_s a_i'(0) - \tau_s\alpha_{is}a_i(0) + q_s\sum_{h=1}^n \omega_{hs}'(0)EY_hY_i - \rho_s\tau_s c_i.$$

In terms of matrix notation, we can write:

$$q_s\Sigma\omega_s'(0) = \tau_s[\rho_s c + a'(0)] + \tau_s\alpha_{\bullet s}\odot a(0).$$

This gives the value for $q_s\omega_s'(0)$ given in Theorem 2.1 (taking into account Theorem 1.2 for $q_s\omega_s(0)$).

C Proof of Theorem 3.1

From expression (B.1), we derive:

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

Moreover, we already showed that:

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

and from the expression of $k''_{is}(\sigma)$:

$$k'''_{is}(\sigma) = 3\alpha_{is}(2a'_i(\sigma) + \sigma a''_i(\sigma)) + (1 + \sigma\alpha_{is})(3a''_{is}(\sigma) + \sigma\alpha'''_{is}(\sigma))$$

and thus: $k'''_{is}(0) = 3a''_i(0) + 6\alpha_{is}a'_i(0)$. Therefore, after replacement in $\varphi'''_{is}(\sigma)$, we have:

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

Recall that:

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

Therefore, after division by $u''_s(q_sR_f)$, we obtain:

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

From the expressions (B.3) and (B.2) pertaining to the second derivatives of W_s and R^s_i we have:

$$\begin{aligned}
\frac{d^3W(\sigma)}{d\sigma^3} \quad = \quad & q_s\sum_{i=1}^{n}\omega'''_{is}(\sigma)\left[R^s_i(\sigma) - R_f\right] + 3q_s\sum_{i=1}^{n}\omega''_{is}(\sigma)\frac{dR^s_i(\sigma)}{d\sigma} \\
& +3q_s\sum_{i=1}^{n}\omega'_{is}(\sigma)\frac{d^2R^s_i(\sigma)}{d\sigma^2} + q_s\sum_{i=1}^{n}\omega_{is}(\sigma)\frac{d^3R^s_i(\sigma)}{d\sigma^3}
\end{aligned}$$

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

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

and therefore:

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

Therefore:

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

This can be written as:

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

However, we already know that:

$$\begin{aligned} \omega_s(0) &= \frac{\tau_s}{q_s}\Sigma^{-1}a(0) = \frac{\tau_s}{q_s}\varsigma \\ \omega_s'(0) &= \frac{\tau_s}{q_s}\Sigma^{-1}\left[\rho_s c + a'(0) + \alpha_{\bullet s}\odot a(0)\right] = \frac{\tau_s}{q_s}\left[\rho_s\varsigma_{[3]} + \Sigma^{-1}a'(0) + \Sigma^{-1}\varsigma_b^{(s)}\right] \end{aligned}$$

After replacement, we obtain:

$$\begin{aligned} \varphi_{is}^{''}(0) = 0 \quad i = 1, ..., n \quad \Leftrightarrow \quad 0 &= \frac{\kappa_s\tau_s}{q_s}d - 4\rho_s\left[\varsigma^\intercal\Gamma_i\omega_s^{'}(0)\right]_{1\le i\le n} \\ &-6\frac{\tau_s}{q_s}(3\rho_s - 1)(\varsigma^\intercal\Sigma\varsigma)\Sigma\varsigma \\ &+\Sigma\omega_s^{''}(0) - \frac{\tau_s}{q_s}\left[a^{''}(0) + 2\alpha_{\bullet s}\odot a^{'}(0)\right] \end{aligned}$$

Replacing $\omega_s^{'}(0)$ by its value above we obtain

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

which is the result of Theorem 3.1.

D 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:

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

with $\varsigma_b = \Sigma^{-1}[\alpha\odot a(0)]$. Hence, the market clearing condition $\sum_{s=1}^{s} q_s\omega_s''(0) = 0$ is tantamount to:

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

We therefore obtain:

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

We now replace $a'(0)$ in the expression above, yielding:

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

E Proof of Theorem 3.3

The value of $\omega_s(0)$ and $\omega_s^{'}(0)$ in Theorem 3.3 are easily deduced from Theorem 2.2 by applying Assumption 3.1. From this assumption and Theorem 3.1 we have:

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

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

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

We now replace $a^{'}(0) = -\overline{\rho}\Sigma\varsigma_{[3]} - \overline{\lambda}\alpha \odot a(0)$ in the expression above to deduce

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

From Theorem 3.1 and Assumption 3.1:

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

We now replace $a(0) = \Sigma\varsigma$, $a^{'}(0) = -\overline{\rho}\Sigma\varsigma_{[3]} - \overline{\lambda}\alpha \odot a(0)$ and:

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

in the expression above to deduce

$$\begin{aligned} \frac{1}{2}\omega_s^{''}(0) = \quad & \tfrac{\tau_s(\overline{\kappa}-\kappa_s)}{2q_s}\Sigma^{-1}d + 3\tfrac{\tau_s}{q_s}\left(\rho_s - \overline{\rho}\right)\left(\varsigma^{\intercal}\Sigma\varsigma\right)\varsigma + \tfrac{\tau_s}{q_s}\overline{\rho}\left(\overline{\lambda} - \lambda_s\right)\Sigma^{-1}\left(\alpha \odot c\right) + \tfrac{\tau_s}{q_s}\overline{\lambda}\left(\overline{\lambda} - \lambda_s\right)\Sigma^{-1}\alpha \odot \left(\alpha \odot \Sigma\varsigma\right) \\ & +\tfrac{2\tau_s}{q_s}\Sigma^{-1}\left[\left(\rho_s\left(\rho_s - \overline{\rho}\right) - \left(\overline{\rho^2} - \overline{\rho}^2\right)\right)\varsigma^{\intercal}\Gamma_i\varsigma_{[3]}\right]_{1\leq i\leq n} + 2\tfrac{\tau_s}{q_s}\Sigma^{-1}\left[\left(\rho_s\left(\lambda_s - \overline{\lambda}\right) - \left(\overline{\rho\lambda} - \overline{\rho}\overline{\lambda}\right)\right)\varsigma^{\intercal}\Gamma_i\varsigma_b\right]_{1\leq i\leq n} \end{aligned}$$

Table 1: Description of Model Parameters, Return Processes and Mutual Funds

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. (6)	Kurtosis-tolerance coefficient
$\sigma\alpha_{is}$	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
$\overline{\rho^2}$ - $\overline{\rho}^2$	Eq. (26)	Dispersion of skewness parameters
$\overline{\lambda\rho}$ - $\overline{\lambda}\overline{\rho}$	Eq. (26)	Covariance of skewness preference parameters and beliefs
ς	Eq. (7)	$\Sigma^{-1}a(0)$ Mean-variance mutual fund
ς_b	Th. 3.3	Beliefs portfolio
$\varsigma_{[3]}$	Th. 3.3	Skewness portfolio
$\varsigma_{[4]}$	Th. 3.3	Kurtosis portfolio
$\varsigma_{b[3]}$	Th. 3.3	Beliefs-about-skewness portfolio
$\varsigma_{[b,b]}$	Th. 3.3	Beliefs-about-beliefs portfolio
$\varsigma_{[3,3]}$	Th. 3.3	Cross-co-skewness portfolio
$\varsigma_{[3,b]}$	Th. 3.3	Cross-co-skewness beliefs portfolio
$\alpha_{\bullet s}$	Eq. (13)	Belief loadings investor s