



HAL
open science

Commodity Markets through the business cycle

Mathieu Gatumel, Florian Ielpo

► **To cite this version:**

Mathieu Gatumel, Florian Ielpo. Commodity Markets through the business cycle. *Quantitative Finance*, 2014, 14 (9), pp.1597-1618. 10.2139/ssrn.2334180 . hal-01302479

HAL Id: hal-01302479

<https://hal.science/hal-01302479>

Submitted on 30 Mar 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Commodity markets through the business cycle

Mathieu Gatumel[‡]

Université Savoie Mont Blanc
IREGE

Florian Ielpo^{‡‡}

Centre d'Economie de la Sorbonne
IPAG Business School

April 27, 2015

Abstract

We introduce a new measure of risk appetite in financial markets, based on the cross sectional behavior of excess returns. Turning them into probabilities through a Markov Switching model, we define one global risk appetite measure as the cross-sectional average of the individual probabilities for each asset to be in a “risk appetite” regime. Given the probabilistic approach that comes naturally with this Markov Switching framework, we present various tests to gauge the interest of the risk appetite measure that is presented here. Using these tests we show that our index behaves well vs. various competitors, especially in out-of-sample results. We test for the information content of various assets and find that a core of asset allocation-related assets provide the best possible choice over various competing specifications.

JEL Classification: E44, F37, G11, G12.

Keywords: Risk Appetite, Cross Market Linkages, Markov Switching, Financial Crisis Prediction.

[‡]Corresponding author. Université Savoie Mont Blanc. 4, Chemin de Bellevue, BP 80439, 74944 Annecy le Vieux CEDEX. Email: mathieu.gatumel@univ-smb.fr.

^{‡‡}University Paris 1 Panthéon - CERMSEM, Maison des Sciences Economiques, 106 Boulevard de l'Hopital, 75013 Paris, France. Email: florian.ielpo@ensae.org

1 Introduction

The evolution of financial asset prices reflects the variations in investors' willingness to bear various types of risks: a greater aggregated appetite for risky assets should lead to an increased demand for these assets and eventually increased risky asset prices. Both academic and practitioners now refer to this phenomenon as "risk appetite". This concept is different from risk aversion and risk itself. Risk aversion refers to a very specific concept in the academic literature, related to the preferences of investors – as summarized into a parameter characterizing the utility function of the representative agent for example – and being regarded as a rather stable parameter as it reflects the psychological ability of a person to bear risk (see [Gai and Vause \(2004\)](#) on this point). It should therefore go through limited times variations, at least from a weekly to monthly perspective. Risk by itself is supposed to reflect the "objective" level of potential drawdowns of financial assets: it is empirically found to be time varying and cyclical. Risk appetite reflects the market participants' evaluation of the potential evolutions of risk, being thus a subjective assessment of potential investment opportunities – and threats: as explained in [Gai and Vause \(2004\)](#), risk appetite reflects the combination of risk aversion and of the macroeconomic environment.

Several attempts to measure risk appetite have been proposed in the literature. These attempts could be gathered into five different categories: a first stream of literature uses principal component analysis to extract the first components implicit in asset returns and uses this measure to gauge risk appetite, as illustrated in [Sløk and Kennedy \(2004\)](#). This approach assumes that risk appetite is the main driver of financial markets, once idiosyncratic risk has been averaged out. A second approach combines the dynamic of returns and the coincident dynamics of risk metrics, such as volatility: [Kumar and Persaud \(2002\)](#) and [Misina \(2003, 2008\)](#) propose to build a risk appetite index based on this joint behaviour of expected returns and risk. A third approach aims at combining different market indicators to proxy risk appetite, such as the Volatility IndeX (VIX), credit spreads or the returns on precious metals such as Gold. [Illing and Aaron \(2005\)](#) and [Coudert and Gex \(2008\)](#) survey various practitioners' measures using such information. A fourth approach aims at uncovering risk appetite from the joint behaviour of asset returns and option prices, as presented in [Bliss and Panigirtzoglou \(2004\)](#) and in [Gai and Vause \(2004\)](#). This approach is based on the intuition that options are hedging investment vehicles and that any surge in risk appetite should lead to a decrease in the valuation of these assets consistently with rising returns on the underlying asset. A last approach proposes structural models combining economic and market data to estimate risk appetite or financial markets' optimism, as presented in [Bekaert et al. \(2009\)](#).

Event though risk appetite is an unobservable phenomenon, these various approaches cast light over the topic of interest here. Most of these articles conclude with the following stylized facts: risk appetite is time varying and persistent. Risk appetite increases with risky assets' returns and decreases with equity volatility. Despite the interest of these findings, the previous approaches have common pitfalls: some of them use ad-hoc data series to measure risk appetite, combining spreads, volatility and returns together in spite of the different nature of these data series. Others cannot be used to build a global risk appetite index, as risk appetite can be regarded as a cross-asset and cross-market phenomenon: for example, options are not available for every market. Finally, given that risk aversion is non-measurable, the "quality" of its measure requires a probabilistic framework that would make it possible to perform econo-

metric tests: this seems to be hardly possible for most of the previously mentioned approaches.

This article aims at dealing with these shortcomings by providing a framework making it possible to combine different kinds of assets to measure risk appetite in a cross-asset perspective. The intuition used here is simple: we assume that risk appetite is the main driver to asset allocation as highlighted in [Sløk and Kennedy \(2004\)](#) and [Coudert and Gex \(2008\)](#) (see Table (1) and the empirical tests showing how statistically strong is the PCA’s first factor for risk appetite). When risk appetite is high, investors have an increased tendency to replace less risky assets by riskier ones in their portfolios. On the contrary, when risk appetite decreases, investors would sell the riskiest assets in their portfolios and replace them by riskless ones, such as government bonds. This type of phenomenon is usually referred to as respectively “flight-from-quality” and “flight-to-quality”, as documented in [Bekaert and Wu \(2000\)](#) and [Bekaert et al. \(2009\)](#): we assume that this phenomenon is what characterizes risk appetite and its swings. When risk appetite is high, a balanced portfolio of risk premia should have a positive return. On the contrary, when risk appetite is undermined, the returns of such a portfolio should have a tendency to be negative. We propose to use this empirical fact to build and gauge our measure of risk appetite, jointly with the previously mentioned stylized facts.

Given the need for a probabilistic approach to build statistical tests, we build our risk appetite index using Markov Switching models, as originally introduced in [Hamilton \(1989\)](#). Using a basket of excess returns, we estimate for each asset the probability for each date in the sample that this asset is in a “risk appetite mode”, that is that it stands more chance to have a lower volatility and a higher expected return. For a given date, the risk appetite index is then the naive average across assets of those probabilities: by doing so, we do not need to assume that the cross asset relations are stable, gaining from the classical effects detailed in the model averaging literature (see e.g. [Claeskens and Hjort \(2008\)](#)). This basket is designed to reflect different risks: it contains both investment grade and high yield credit risk, as well as large and small capitalization equity risks for the US and for the Euro zone. By using such a sample, we will be able to discuss risk aversion from a cross-asset and a global perspective using the most natural prism to discuss risk appetite: an asset allocation perspective. To gauge the quality of our approach, we use two different criterions: we first use a financial one that is the profit-and-loss associated to a strategy that would be long or short this basket depending on the measured risk appetite. When our risk appetite measure would be accurate, such a strategy would deliver a positive return, outperforming for example a long-only position. On top of this criterion, we add a statistical one: given that our risk appetite index is built from time series models, we are able to perform density-based tests, such as density forecasts as presented in [Diebold and Mariano \(1995\)](#) and [Amisano and Giacomini \(2007\)](#). Both tests are performed in- and out-of-sample.

We provide various tests to discuss the accuracy of our modelling approach, comparing it to different other modelling choices and globally find that this approach consistently yields the best results. Then, given the fact the Markov Switching model turns anything into probabilities – that are a comparable metrics – we test whether adding asset classes such as foreign exchange rates, emerging equities or commodities improves the measurement of risk appetite. We also discuss in this perspective the inclusion of the VIX index and of credit spreads in a risk ap-

petite measure. Our results show that our risk appetite measure behaves fairly well, matching the key characteristics one would expect from such a measure, mainly through out-of-sample metrics. We then test whether additional dataset would have any additional information regarding the variation of risk appetite: we find that taken as groups such as “commodities” there is no extra information. Taken individually, we find that selected emerging equities, the Sterling, the Yen, the 10-year US and German rates or the VIX taken individually stand a chance to improve the RAI, mainly in an in-sample manner. We also test the informational content of various economic series. For most of the series investigated here, we fail at finding additional information when using an econometric test. Still, we find an economic value to these data in- and out-of-sample, when considered all together.

This paper is organized as follows: Section 2 discusses the building of our risk appetite index and of the competing approaches to which it will be compared. Section 3 presents the various metrics used to compare models and datasets. Section 4 presents the dataset used in the article. Section 5 presents the results of our main approach, comparing it to other possible choices. Finally, Section 6 presents results regarding the opportunity of adding additional dataseries to our risk appetite measure.

2 Construction of the Risk Appetite Index

This section aims at presenting the construction of our risk appetite index. Let $r_t^{(i)}$ be the excess return on asset i at time t , that is the return on asset i in excess of the return on the risk free asset. The total number of asset is I , that is that $i = 1, \dots, I$. Now, we assume that each asset is characterized by two different kinds of regimes: one of these regimes is a flight-to-quality regime, whereas the other one is a flight-from-quality one. We thus have:

$$r_t^{(i)} \sim N(\mu_j^{(i)}, \sigma_j^{(i)}), \text{ with } s_t^{(i)} = j, \quad (1)$$

when state j prevails. Given that we focus on a two-regime model, $j = 1, 2$. The description of a flight to/from quality is done through the following restrictions:

$$\mu_1^{(i)} > \mu_2^{(i)} \quad (2)$$

$$\sigma_1^{(i)} < \sigma_2^{(i)}. \quad (3)$$

Thus, regime 1 expresses a higher degree of risk appetite than regime 2: risk appetite is thus assumed to be an increasing function of expected returns and a decreasing function of risk. This property is consistent with the assumption of [Misina \(2003, 2008\)](#) and is essential to any risk appetite measure.

Now, building up on the previous notations, we can write the probability that state 1 prevails at time t as follows:

$$P(s_t^{(i)} = 1) = P(s_{t-1}^{(i)} = 1) \times p_{11}^{(i)} + P(s_{t-1}^{(i)} = 2) \times p_{21}^{(i)}, \quad (4)$$

where $p_{jk}^{(i)}$ are time invariant transition probabilities, and parameters to this switching model. Once these parameters have been estimated by maximum likelihood as presented in Hamilton

(1989), we are able to compute the following index for a given date t :

$$\pi_t = \frac{1}{I} \sum_{i=1}^I P(s_t^{(i)} = 1). \quad (5)$$

We label this simple average of the flight-from-quality probabilities “risk appetite index” for the date t . Given the way this risk appetite measure is computed, we will refer to it as “the model averaging approach” in the rest of the article.

Beyond this apparent simplicity, this index has three interesting properties:

1. The estimation of a 2-regime Markov Switching model using univariate series has a lower numerical cost than estimating a multivariate switching model, decreasing the parameter uncertainty and thus increasing the accuracy of the measure.
2. Transforming returns into probabilities turns asset returns that are not always comparable – in terms of scales or variance for example – into comparable quantities. We present here the computation of the risk appetite indicator based on a dataset of returns. However, instead of returns, we could use the VIX or credit spreads variations to estimate these probabilities. On top of that, the average of these probabilities is again a probability: it is thus easily interpretable in terms of scale, unlike most of the measures presented in the existing literature.
3. Finally, given that we use a simple average – which is a rather naive choice – we do not assume that risk appetite always obeys to the same sequence of events: any of the assets i can be the source of a sudden drop in risk appetite. This feature would be lost in a multivariate Markov switching model.

To this index, we add three competitors with a greater amount of rigidity than the Risk Appetite Index (RAI):

1. A multivariate Markov Switching with two regimes again: regime 1 is the risk appetite regime. With this approach, risk appetite is measured by the probability at time t that assets $i = 1, \dots, I$ are *jointly* in regime 1. Such a measure is more difficult to estimate given that I can potentially be high. On top of that, given the rigidities of the multivariate MS(2) model, it implies that strong drops in risk appetite always occur the same way, starting from the same assets and then contaminating the other ones through a similar sequence.
2. An univariate MS(2) model estimated from the risk appetite index of [Kumar and Persaud \(2002\)](#) and from the risk appetite index of [Misina \(2008\)](#). For these competitors, we start by computing the risk appetite index as presented in both these references. Then, to be able to compare risk appetite measures through similar metrics, we turn each of them into probabilities using a Markov switching model with two regimes.

3. An univariate MS(2) model estimated from the returns on the first factor of a PCA analysis obtained from the covariance matrix of the returns of the assets $i = 1, \dots, I$. The first factor of the PCA is usually regarded as a “market” portfolio, that is a portfolio that best synthesizes the evolution of the underlying assets. When computed from excess returns, this portfolio – when hold by an investor – turns out to deliver positive performances over “risk-on” periods, that is when the valuation of risky assets rises. Hence, when estimating a Markov switching model with two regimes, similar conditions to Equations (2) and (3) are very to be found. As attractive as this setting looks, the estimated covariance matrix necessary to compute the PCA factors introduces rigidities that are likely to diminish the interest of this approach when compared to our risk appetite measure.

A thorough comparison of the RAI and these three competitors will cast light on the dynamics of risk appetite in financial markets.

3 Comparison methodologies

To gauge the ability of our indicator to measure risk appetite, we use two different kinds of metrics. The first one is a financial metric: when our indicator would make it possible to measure variations in risk appetite, it would then make it possible to obtain a positive return over a core dataset of excess returns, timing the variations of risk taking in financial markets. The returns on an investment strategy based on this indicator will be thus one of our key quality measures. We refer to that metric through the label “economic value”. Second, we also use an econometric measure, gauging the ability of our risk appetite measure to capture risk appetite in a statistical way. Using a density based test, we provide both in- and out-of-sample results regarding the goodness of fit of the proposed model using various kinds of excess returns datasets. In this section, we present the two test methodologies. For both of them, we thus present results in terms of coincident measure and in terms of forecasts: this is related to the stylized fact observed in the previously listed literature that risk appetite is a *persistent* phenomenon. A proper measure for it should have in some way a forecasting power over it, that is related to the accuracy of its persistence modelling.

3.1 Economic Value

We compare our risk appetite indicators based on their economic value. When a given risk appetite measure is supposed to provide an accurate measure of this phenomenon, then it should be able to deliver a positive return when using it as an investment signal for the dataset of returns $r_{t+k} = (r_{t+k}^{(1)}, r_{t+k}^{(2)}, \dots, r_{t+k}^{(I)})$. We compute this economic value using the following two steps:

1. First, we create a portfolio out of the returns that have a high correlation with risk appetite. This factor is built as the first factor of a principal component analysis obtained using the sample of returns $r_{t+k} = (r_{t+k}^{(1)}, r_{t+k}^{(2)}, \dots, r_{t+k}^{(I)})$: when performing this PCA, we use the first eigenvector as portfolio loadings. We denote \tilde{r}_t the return on this portfolio at date t . Here, we assume that this portfolio is positively related to risk appetite: the

higher the risk appetite and the higher is the value of this portfolio¹.

2. Then, using the risk appetite indicator, we compute a profit-and-loss obtained when buying the previous portfolio when $\pi_t > 0.5$ and selling it when $\pi_t < 0.5$. This approach in terms of probabilities leads us to a natural choice in terms of threshold triggering long and short positions. These Profit and losses will help us decide which model provides us with the best risk appetite indicator.

We compute the returns over such investment strategies for two different cases: we compute our index both using the current probability to be in a risk appetite regime and the forecast probabilities to be in such a regime. Again risk appetite is supposed to be a persistent phenomenon, which is why we work both with our coincident measure and with forecasts of risk aversion. These forecast probabilities are obtained by iterating Equation (4). We compute these probabilities both in the in- and out-of-sample cases.

3.2 Econometric comparison

We also compare competitors through their ability to capture the empirical distribution of the excess returns of assets that we assume to be related to risk appetite. Let $f_{M_a,k}(r_{t+k}|\hat{\theta}_{M_a})$ be the estimated density for model M_a with a forecast horizon equal to k , with $\hat{\theta}_{M_a}$ the estimated parameters and with $r_{t+k} = (r_{t+k}^{(1)}, r_{t+k}^{(2)}, \dots, r_{t+k}^{(I)})$. When $k = 0$, $f_{M_a,k}(\cdot)$ is thus the in-sample density of the set of returns. When $k > 0$, $f_{M_a,k}(\cdot)$ is a forecast density, that we compute in an out-of-sample manner.

A proper metrics of the ability of a given time series model to capture the relevant empirical features of a given set of data is the loglikelihood associated to the model investigated:

$$L_{M_a,k} = \frac{1}{T - \tau - k_{max} + 1} \sum_{t=\tau+k_{max}-k}^{T-k} \log f_{M_a,k}(r_{t+k}|\hat{\theta}_{M_a}), \quad (6)$$

where k is the forecast horizon, τ is the number of observations used for the first in-sample window, T is the total number of observations and k_{max} is equal to the maximum forecast horizon that will be used in our empirical computations. Again, the coincident analysis is done with $k = 0$.

In order to compare two different models – say model M_a and model M_b – we use the following statistics:

$$t_{M_a,M_b}^k = \frac{L_{M_a,k} - L_{M_b,k}}{\frac{\hat{\sigma}_{M_a,M_b}}{\sqrt{T-\tau-k+1}}}, \quad (7)$$

with $\hat{\sigma}_{M_a,M_b} = \sqrt{V[L_{M_a,k} - L_{M_b,k}]}$. This variance is estimated using a Newey-West estimator. Under the null hypothesis that both modelling approaches deliver similar performances –

¹The intuition behind it is related to [Sløk and Kennedy \(2004\)](#): the main risk factor in an asset allocation portfolio is considered to be risk appetite, that is a factor to which all risky assets are strongly related. A PCA makes it possible to estimate this factor, as shown in [Table \(2\)](#). However, it can hardly be regarded as a risk appetite measure given that it consists of volatile returns that does not thus exhibit the required persistence.

would it be in terms of forecasts when $k > 0$ or in terms of in-sample fit when $k = 0 - t_{M_a, M_b}^k$ follows a standard Normal distribution. A significant positive (negative) estimated value rejects the null of equal performance between competing forecasts, and provides evidence in favor of model M_a (M_b).

When $k = 0$ this test is exactly the density test proposed by [Vuong \(1989\)](#). When $k > 0$, this test amounts to the [Diebold and Mariano \(1995\)](#)'s density forecast test.

We increase our understanding of both the in-sample fit and the density forecasts by breaking down these tests into subtests using the weights functions proposed in [Amisano and Giacomini \(2007\)](#). We change Equation (6) the following way:

$$L_{M_a, k} = \frac{1}{T - \tau - k_{max} + 1} \sum_{t=\tau+k_{max}-k}^{T-k} \omega(r_{t+k}) \log f_{M_a, k} \left(r_{t+k} | \hat{\theta}_{M_a} \right), \quad (8)$$

where $\omega(r_{t+k})$ is one of the following functions, allowing us to focus on different pieces of the support of the distribution of the returns:

1. $\omega_0(r_{t+k}^{st}) = 1$,
2. $\omega_3(r_{t+k}^{st}) = \Phi(r_{t+k}^{st})$, where $\Phi(\cdot)$ is the standard cumulated normal distribution function,
3. $\omega_4(r_{t+k}^{st}) = 1 - \Phi(r_{t+k}^{st})$, where $\Phi(\cdot)$ is the standard cumulated normal distribution function,

where r_{t+k}^{st} is the vectors of the standardized returns. Using such a breakdown of the returns' distribution's support will help us understand under which circumstances which model performs better.

4 Dataset

The dataset used here is made of different asset prices, spreads or volatilities. Our dataset is split into three types of datasets:

- A core dataset, building on the intuition of the flight-from/to-quality periods applied to the most basic asset allocation assets. These assets are the following:
 - Equities: we consider different types of equity indices for the US and the EMU investment universe. First, we consider both large and small capitalization indices for the developed world stock market. In the US case, the large caps index is chosen to be the SP500 and the small caps one is the Russel 2000. In the EMU case we use Eurostoxx 50 and the MSCI small caps EMU.
 - Credit: Investment grade and high yield bond indices for both the US and the EMU are considered. These indices are Bank of America Merrill Lynch's indices.

These “core assets” are turned into excess returns to reflect to gain/cost of the opportunity to hold these risky assets vs. riskless ones. The return on a riskless investment is proxied through Merrill Lynch’s government bond indices – US and EMU separately – whose returns are used to compute these excess returns.

- “Non core assets” which are asset that can potentially be affected by risk appetite, but that are not included traditionally in the most basic asset allocation of traditional asset managers:
 - Emerging equities: we focus our attention on broad indices that are segregated depending on geographic arguments: we use the MSCI EM Asia, Latam and Europe.
 - Foreign Exchange rates: an important part of the financial transactions around the world come from the FX investment universe. We retained seven exchange rates for their ample liquidity and well known economic interest: the Euro, the Swiss Franc, the British Pound, the Norwegian Krone, the Yen, the Brazilian Real and the Singapore Dollar against the Dollar.
 - Commodities: we use five additional series of returns from the commodity universe. We focus on the Dow Jones broad commodity index and on four specific indices: GSCI Agriculture, Energy, Industrial Metals and Precious Metals. The first index is a diversified index representing the whole commodity markets, as it includes both hard and soft commodities.
- To these series we add financial series that are traditionally used into most of the practitioners’ risk appetite indices:
 - the Volatility Index (VIX) and the VDAX index, that are option-implied measures of volatility respectively computed out of options whose underlying assets are the SP500 and the German DAX. .
 - Investment Grade and High Yield option adjusted spreads, both in the US and in the German case obtained from the Merrill Lynch dataset.
 - The returns on Gold, as Gold is usually considered as a safe haven from a cross asset perspective.

The dataset starts on January 8th of 1999 and ends on December 12th of 2011. This period was selected as it stands a good chance of being as stationary as possible: this period is at the end of the disinflationary period that starts in 1979 with the Volcker era – which is essential in terms of bonds’ behavior – and it covers the moment when China joined the World Trade Organization – which is likely to have a tremendous impact in terms of emerging world

related assets. The data frequency is weekly: as we focus on regimes, we need to find enough persistence in assets' returns to obtain reliable estimates of the MS parameters. The weekly frequency offers a balanced mix between non Gaussian returns and a stronger persistence than daily data.

Descriptive statistics are presented in Table (1). Table (2) presents the correlation of each of these data returns or variations with the first factor of a principal component analysis performed over the core assets. Several conclusions arise: first of all, over the total sample, the correlation of each asset (excepted for the Swiss Franc against US Dollar) with the factor is statistically different from zero. When examining the R2 obtained from a regression of each asset's returns on the first factor of the PCA analysis performed over the core assets, we get the intuition that this factor explains a lot of the financial market's behavior. For example, this factor explains 84% of the excess returns on the SP500 and 79% of that of the Eurostoxx case. It also explains 42% of the VIX and VDAX variations, 22% of the US 10-year rate variations and 46% of Latam equities' returns. Its explanatory power seems to be weaker in the case of currencies and commodities – even though most of their long term correlations to this factor are found to be statistically different from zero. Beyond this explanatory power, core assets, emerging equities, commodities and rates are positively related to this factor, while volatility indices are negatively related to them. When it comes to currencies, the picture is mixed: USDBRL is negatively related correlated to this factor, whereas USDJPY is positively correlated to it. This has to be related to the position of each currency as a relative safe haven: USD is a safe haven relatively to the Brazilian Real but not to the Japanese Yen – as our results suggest it.

When slicing and dicing the sample into subsamples, we obtain a pretty similar picture when it comes to the sign of these correlations – but in the case of certain exchange rates as the Euro or the Swiss Franc vs. Dollar. The main difference between each subsample is clearly the intensity of the relation between these assets and the PCA factor. For example, the correlation between the US High Yield and this factor over 1999-2004 is equal to 70% whereas it is equal to 55% over 1999-2004. This is consistent with the degree of freedom that our RAI will provide us with.

5 Economic and econometrics comparisons of the RAI and its competitors

In this section, we review the results obtained when estimating the RAI and when we compare it to its competitors.

5.1 Estimation results

Before comparing to its competitors, we first discuss the estimation results of the risk appetite index described in Equation (5).

Table (3) presents the expectations and volatilities for each asset estimated from the risk aversion index obtained using a method of moment estimator² over our core dataset. As the table

²We compute sample averages for each asset depending on the RAI as computed from Equation (5): when

makes it clear, regime 1 is consistently characterized by higher returns and lower volatilities – which is one of the main features of a risk appetite regime – than regime 2. For example, in the SP500 case, when the average excess return in regime 1 is estimated to be -2.211%, for regime 2 this figure is now -22.287%. Its volatility in regime 1 (16.953%) is much lower than its volatility in regime 2 (38.927%). This makes our hypothesis in Equations (2)-(3) empirically valid and this conclusion holds for each of the assets in this core dataset.

Figure (1) presents the time series evolution of the risk appetite index over the 1999-2011 period. The RAI is compared to the NBER crisis periods and selected key dates of the financial history. The RAI follows globally speaking the key dates and matches the recession periods: it drops at the beginning of the 2001 crisis period and starts dropping from mid 2007 with the start of the US housing crisis. Interestingly, the RAI points out two periods for which the drop in risk appetite is not coincident with NBER economic crisis as during the EU debt crisis case in 2011. The difference with such periods is that they seem to have a shorter duration than those occurring during a NBER crisis (see for example the 2008 case). Figure (3) compares the RAI to the cumulated first factor of the PCA: the graph clearly shows how periods over which the RAI is below 50% are coincident with periods of drops in the cumulated first factor, casting additional light on the tight link between risk appetite and the returns in our core dataset.

Figure (4) compares the model averaging approach to the other computation methods. Both multivariate and PCA approaches agree on the periods of drops in risk appetite. Still, the model averaging approach exhibits a higher degree of persistence – which is an important feature of risk appetite – motivating the forthcoming economic and econometric comparison between these various measures. On the contrary, correlation based methods do not clearly highlight the periods for which investors present risk appetite. In addition to that, they are not characterized by a high degree of persistence.

5.2 Model comparison

We compare the model averaging approach proposed in this article to four competing specifications: the traditional risk aversion index of Kumar and Persaud, the approach of Misina (2003), a multivariate Markov switching model based risk appetite indicator and a univariate one estimated from the time series of the first PCA factor’s returns, as explained in Section 2.

The in-sample results are presented in Tables (6) and (7). Table (6) presents the in-sample econometric comparison. From these tests results, we find that the proposed RAI is only dominated by the multivariate Markov Switching approach over the full period considered here, with a test statistics equal to -4.49. The RAI is consistently dominated for the subperiods in the table, but for the 2008-2011 one for which both models are considered to be similar³. On the contrary, the RAI consistently dominates the other approaches. For 1999-2004 and

the RAI is greater or equal to 0.5 at time t , we say risk appetite prevails. We compute the conditional moments using this information.

³We do not reestimate the models for these sub periods for our in-sample results: we only compute the log-likelihood test (Vuong (1989)) over different samples so as to evaluate the stability of our results. This remarks hold for the forthcoming tables.

2008-2011, the first factor and the model averaging approach are equivalent. From these results, we conclude that even in an in-sample setting the RAI provides interesting results. It seems however that measuring accurately the transmission of shocks – through the multivariate transition matrix – is of a primary importance here.

Table (7) presents results regarding the economic value of each of the five methods investigated here. Globally speaking, the model averaging approach dominates the rest of the models in terms of returns but for the 1999-2004 and 2005-2007 periods. For this period, the realized returns is equal to 9.79% per year with the RAI whereas it is only equal to 1.378% in the first factor case. Now, the RAI still delivers a lower kurtosis which is an interesting feature. We can also point out that the RAI-KP and RAI-MI approaches are the worst competitors in terms of returns but for the 2005-2007 period.

Table (8) presents the [Amisano and Giacomini \(2007\)](#)'s density forecast test results comparing the five approaches for four forecast horizons, that is one to four weeks ahead forecasts. This table focuses on out-of-sample density forecasts⁴. The global results show that the model averaging approach clearly dominates its competitors but for the 2010-2011 sample over which the multivariate MS model is favored for one week ahead forecasts and found to be equivalent to the other approaches. Beyond this point, the model averaging approach provides consistently better density forecasts across all forecast horizons. Interestingly the test statistics are increasing functions of the forecast horizon: it seems that the longer the horizon and the stronger is the model averaging relative advantage over its competitors.

Table (9) presents the out-of-sample results for the economic value comparison. The conclusions obtained from this table are quite similar to that of the previous one: the model averaging approach leads to the highest returns but for the four weeks ahead forecasts for which the multivariate model dominates. The PCA, RAI-KP and RAI-MI models are always discarded as a possibility. For example in the one week ahead forecast case, the model averaging approach delivers an annualized return of 5.649% for a risk budget of 10%, whereas the first factor model's return is -3.621%, the multivariate model's is 4.0.69%, the RAI-KP is -0.719% and the RAI-MI is 3.827%.

A least and natural question, coming with our risk appetite index is the level to be reached by the index for a period to be labeled as a "risk appetite" one. Given that our index is a probability, 50% is a good candidate. However, given the test methodologies that come with our approach, we ran tests to decide whether 50% is the right threshold or not. Those tests are presented in Tables (4) and (5). Both from an economic and econometric point of view, the 50% threshold either delivers comparable or superior results to the other potential thresholds investigated, ranging from 0.2 to 0.8. Given those findings, for the rest of the article, we rely on an index using the 50% threshold.

From now on, we only focus on the model averaging approach itself using the 50% threshold as the previous' section showed how it has an interesting in- and out-of-sample behavior.

⁴These out-of-sample density forecasts are obtained through a rolling estimation scheme, with an increasing window size. The first forecast uses the 1999-2007 period as an estimation dataset.

5.3 The information content of additional data

One key feature of the approach developed here is that it makes it possible to combine any type of data together by first turning these data into probabilities. We intend to make use of this feature to test whether different types of financial assets and different types of macroeconomic data contain implied information about the measure and the forecast of global risk appetite (or not).

5.3.1 The information content of additional assets

These series have been listed previously in Section 4 and cover six types of assets: emerging equities, currencies, commodities, rates, credit spreads and implied volatility measures. Our empirical strategy unfolds as follows: we test whether these series contain information regarding the flight-to/from-quality phases. This means that when it comes to measuring the marginal economic or econometric value of these assets, we do not need to add them to our core asset for our density or economic value test: we simply turn these series into probabilities that we add to the RAI as defined in Equation (5). Hence we have two types of datasets. First, the classical asset allocation one used previously and made of excess returns on equity and credit that spans the symptom of risk appetite : flight from quality. this first dataset will thus consistently be used to gauge the economic and macroeconomic quality of our risk appetite measure. Next, a dataset used to create the risk appetite index itself, that will be based on both the core dataset and the potentially relevant non traditional assets that are about to be investigated.

This equation can be thus rewritten as follows:

$$\pi_t = \frac{1}{I_c + I_a} \left\{ \sum_{i=1}^{I_c} P(s_t^{(i)} = 1) + \sum_{j=1}^{I_a} P(s_t^{(j)} = 1) \right\}, \quad (9)$$

where I_c is the number of core assets and I_a the number of additional assets used to build the RAI. We then compare again the results obtained with the previous metrics. Again, only the RAI is modified here: the rest of the computations are done using the core dataset as it is the only dataset that relates directly to risk appetite, as Table (2) shows.

Table (10) presents aggregated results for this marginal analysis, and Figure (5) compares the RAI to its variations depending upon the various dataset extensions examined here. The in-sample results show that every group of data has a negative econometric contribution, without necessarily leading to a decreased economic value. This is the case for the rates, with a density test equal to 7.874 but an extra return of 25 basis points. When it comes to individual assets, we get a more mixed picture. The data series with an economic interest are the Swiss Franc, the Singapore Dollar, the DJ Commodity, the GSCP precious metals, Gold, the US 10 years rates, the German 10 years rates, the spreads (excepted the High Yield) and the VDAX. Despite these in-sample results, most of the series investigated here do not improve the risk appetite measure, would it be in terms of economic value or in terms of density forecast, when looking at the out-of-sample results. This is especially true when considering groups of series. For example, when adding currencies to the construction of the RAI, the

out-of-sample economic value decreases by 1.994% per year whereas the density forecast test is equal to 1.08, making the two datasets comparable at a 10% risk level. There are however a few exceptions that are listed here:

- When considering emerging equities individually, they increase the economic value while being considered as equivalent on a density-based metric. For example, the EM Asia equity index increases the economic value related to the RAI out-of-sample by 1.904%.
- Amongst of currencies, the Swiss Franc, the British Pound, the Brazilian Reais and the Singapore Dollar taken individually would increase the economic value of the RAI without statistically improving the density forecasts: for example, the GBP leads to an improvement of the out-of-sample economic performance by +1.319%.
- In terms of rates and spreads, only the US 10-year rate do not increase again the economic value. All of them do not improve the density forecasts.

Interestingly, the volatility measures that are widely used by practitioners to estimate risk appetite are not found to bring about an improvement of any sort, as both the VIX and the VDAX show similar results, slightly improving the economic value of the RAI. Therefore, such measures are not bringing extra information regarding the swings in risk appetite when compared to the core dataset, suggesting some redundancy in both datasets. Finally, we test specifically the information implied in Gold when it comes to risk appetite measurement as it is usually assumed in financial newspapers that it has a “safe haven” role in asset allocation. Our results shows that Gold has only an in-sample economic appealing contribution, with an incremental return of 12 basis point. However, as for every dataseries in this dataset, this result does not hold out-of-sample. Thus Gold does not help improving the risk appetite forecast.

These results are confirmed by Tables (11) to (14) that display additional forecast horizon on a group by group of series level.

5.3.2 The information content of macroeconomic data

Following the investigation of the additional information contained in other financial data series, we now turn our attention towards economic information. A forecasting power of macroeconomic series over risk appetite is usually assumed by practitioners, but little evidence around their leading properties on risk appetite has been accumulated so far. This section aims at filling this gap, from an empirical point of view.

We use different kinds of data, gathered into three categories: our dataset includes business climate indicators such as the US ISM, the German IFO and the Eurostat Industry Survey. It also comprises employment data, such as the US Non Farm Payroll, the US unemployment rate and the European unemployment rate. Finally, the last type of data is made of inflation-related figures: the US Consumer Price Index and the Eurozone one. With these figures, we intend to cover different types of figures (survey and hard data) coming from two large economic zones (the USA and the Eurozone), over the same period as for the financial assets.

Now, one important difference with our previous investigations is that those figures only come on a monthly basis. Therefore, we cannot directly compare the results obtained previously with those that we are about to discuss. To match this frequency, we turn the dataset of asset prices into a dataset of end-of-the-month closing prices from which we compute monthly returns. We match returns and economic data based on the month of disclosure of each economic figure (and with the month over which each figure has been computed). Again, the period covers 1999-2011, similarly to our previous analyses.

We ran the same tests as before, using this different frequency though. To help the reader understand the main impact of this change in frequency, we chart the risk appetite measures with and without this extra-layer of economic information in Figure (6). Obviously, with this weaker frequency, our monthly measure has a lesser variability, displaying smoother variations when compared to the weekly measure.

Table (15) presents both the economic and the econometric tests' results comparing the incremental information embedded in this economic dataset. As the computations are exactly similar to those in Table 8, we do not explain again here the methodology that we used, but rather focus on the empirical conclusions that can be raised from them. Our in-sample dataset now includes 119 data points, whereas our out-of-sample dataset is made of 36 observations (and thus covers the 2009-2011 period).

From an econometric standpoint, this macroeconomic data does not seem to have any particular kind of extra-information: in- and out-of-sample results show that for most figures, the information content of economic series is equivalent to market data in the out-of-sample case. In-sample, we accept the null hypothesis that economic figures do not bring additional information as well. In the specific case of the US ISM and of the Eurozone employment rate, we find that the version of our risk appetite measure computed without them dominates the version using these additional series. A similar conclusion is reached when adding both inflation series. Now, from an economic point of view, we obtain a more mitigated picture. On the one hand, for 6 individual series out of the 8, we find that economic series have an actual economic value, as they help the strategy generating an additional return that can reach up to around 1% in the ISM case. Out-of-sample, most of these gains vanish. On a group-of-series case, we obtain an increase in the economic value only in the case of CPIs. This additional economic value is consistent with our in- and out-of-sample experiments. Finally, when considering all these variables together, we obtain a global conclusion that is very consistent with the individual results: the density tests show that using economic variables does not result into additional information regarding risk aversion. On the contrary, economic value is clearly increased through the use of economic series, both in- and out-of-sample: in this latter case, the extra return even reaches around 5% per year. One potential explanation for this puzzle is that for a couple of specific events, the economic information helped the investment strategy over the 2009-2011 period. When observing Figure 6, we see that the only difference between the two risk appetite measures appears in 2010, when the one including the economic variables goes above the 0 threshold faster than the one excluding these variables. This period matches a strong pick-up of risky assets that the economic figures based index helps capture faster.

6 Conclusion

This article thus proposes a new measure of risk appetite that entitles its user to perform comparisons and tests in a natural way: given that it is based on a probabilistic framework, our approach makes it possible to compare various assets for the information that they may contain when it comes to risk appetite measurement and prediction. On top of that, it also makes it possible to compare various model specifications. Using these features, our empirical results indicate that the risk appetite index proposed in this article is a robust alternative out-of-sample against other potential competitors. Empirical tests show that a dataset made of asset allocation-related assets provides us with the most reliable measures and forecasts of risk aversion across various alternatives, including economic series.

References

- Amisano, G. and Giacomini, R. (2007). Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business & Economic Statistics*, 25:177–190.
- Baek, I.-M. (2006). Portfolio investment flows to asia and latin america: Pull, push or market sentiment? *Journal of Asian Economics*, 17(2):363–373.
- Baek, I.-M., Bandopadhyaya, A., and Du, C. (2005). Determinants of market-assessed sovereign risk: Economic fundamentals or market risk appetite? *Journal of International Money and Finance*, 24(4):533–548.
- Baker, M. and Wurgler, J. (2007). Investor sentiment in the stock market. NBER Working Papers 13189, National Bureau of Economic Research, Inc.
- Bandopadhyaya, A. and Jones, A. L. (2006). Measuring investor sentiment in equity markets. *Journal of Asset Management*, 7(3):208–215.
- Bekaert, G., Harvey, C. R., and Ng, A. (2005). Market integration and contagion. *The Journal of Business*, 78(1):39–70.
- Bekaert, G., Hoerova, M., and Scheicher, M. (2009). What do asset prices have to say about risk appetite and uncertainty? Working Paper Series 1037, European Central Bank.
- Bekaert, G. and Wu, G. (2000). Asymmetric volatility and risk in equity markets. *Review of Financial Studies*, 13(1):1–42.
- Bliss, R. R. and Panigirtzoglou, N. (2004). Option-implied risk aversion estimates. *Journal of Finance*, 59(1):407–446.
- Bollerslev, T., Gibson, M., and Zhou, H. (2011). Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities. *Journal of Econometrics*, 160(1):235–245.
- Campbell, J. Y. and Cochrane, J. (1999). Force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy*, 107(2):205–251.
- Claeskens, G. and Hjort, N. L. (2008). *Model selection and model averaging*. Cambridge series in statistical and probabilistic mathematics. Cambridge University Press, Cambridge, New York.
- Coudert, V. and Gex, M. (2008). Does risk aversion drive financial crises? testing the predictive power of empirical indicators. *Journal of Empirical Finance*, 15(2):167–184.
- Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3):253–63.
- Gai, P. and Vause, N. (2004). Risk appetite: Concept and measurement. Technical report, Bank of England.
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–84.
- Illing, M. and Aaron, M. (2005). A brief survey of risk-appetite indexes. Technical report, Financial System Review, Bank of Canada.
- Kumar, M. S. and Persaud, A. (2002). Pure contagion and investors’ shifting risk appetite: Analytical issues and empirical evidence. *International Finance*, 5(3):401–36.
- Lemmon, M. and Portniaguina, E. (2006). Consumer confidence and asset prices: Some empirical evidence. *Review of Financial Studies*, 19(4):1499–1529.
- Misina, M. (2003). What does the risk-appetite index measure? Working Papers 03-23, Bank of Canada.
- Misina, M. (2008). Changing investors’ risk appetite: Reality or fiction? *European Journal of Finance*, 14(6):489–501.

Sløk, T. and Kennedy, M. (2004). Factors driving risk premia. OECD Economics Department Working Papers 385, OECD Publishing.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57(2):307–33.

A Tables

A.1 Descriptive Statistics

		Average Return (ann.)	Volatility (ann.)	Sharpe Ratio	Skewness	Kurtosis
Core Assets	Investment Grade US	0.004	0.035	0.11	-2.28	24.96
	High Yield US	0.01	0.099	0.1	-1.39	13.47
	Russell 2000	-0.013	0.274	-0.05	-0.45	2.43
	High Yield EMU	-0.006	0.12	-0.05	-1.23	9.74
	MSCI Small Cap	-0.015	0.244	-0.06	-1.26	5.31
	Investment Grade EMU	-0.003	0.027	-0.11	-1.69	15.57
	SP500	-0.056	0.222	-0.25	-0.46	3.65
	Eurostoxx	-0.075	0.254	-0.3	-0.71	4.21
Other Assets	Gold	0.133	0.187	0.71	0.05	2.66
	GSCI Prec. Metals	0.134	0.198	0.68	-0.39	2.39
	MSCI EM Latam	0.149	0.235	0.63	-0.63	4.75
	Inv. Grade EMU OAS	0.149	0.326	0.46	0.83	3.52
	MSCI EM Europe	0.115	0.299	0.38	-0.2	9.11
	GSCI Ind. Metals	0.092	0.247	0.37	-0.56	2.71
	GSCI Energy	0.109	0.339	0.32	-0.7	2.07
	MSCI EM Asia	0.067	0.228	0.29	-0.47	2.22
	DJ Commodity	0.047	0.181	0.26	-0.89	3.15
	Inv. Grade US OAS	0.062	0.251	0.25	1.08	5.78
	USDBRL	0.018	0.177	0.1	1.19	5.92
	High Yield US OAS	0.028	0.311	0.09	1.04	3.83
	EURUSD	0.009	0.108	0.08	-0.27	0.97
	High Yield EMU OAS	0.01	0.357	0.03	0.71	3.26
	VIX	-0.018	0.893	-0.02	0.51	1.68
	VDAX	-0.032	0.766	-0.04	0.4	2.68
	GBPUSD	-0.005	0.098	-0.05	-0.56	3.39
	GSCI Agri.	-0.026	0.217	-0.12	-0.26	2.03
	USDNOK	-0.018	0.122	-0.15	0.39	0.74
	German 10Y Rate	-0.055	0.217	-0.25	-0.33	3.04
	USDCHF	-0.03	0.116	-0.26	0.38	3.67
	US 10Y Rate	-0.069	0.257	-0.27	-0.04	2.52
	USDJPY	-0.031	0.109	-0.28	-0.2	1.14
	USDSGD	-0.02	0.051	-0.39	0.48	3.11
	US 2Y Rate	-0.221	0.536	-0.41	0.39	4.06
	German 2Y Rate	-0.261	0.564	-0.46	-2.24	21.82

Table 1: Descriptive statistics of the returns on the assets considered in the dataset.

The statistics presented in the table are computed using logarithmic returns over the period that starts on 08/01/1998 and ends on 12/31/2011. The data frequency is weekly. Both the standard deviation and the average returns are scaled into yearly quantities for ease of reading. For Core Assets, the returns are net of government bonds rates.

Asset Class	Asset	R2	1999-2010	1999-2004	2004-2008	2008-2011
US	SP500	0.84	94%*	91%*	94%*	97%*
	Russell 2000	0.83	94%*	92%*	93%*	95%*
	Investment Grade US	0.10	45%*	32%*	35%*	52%*
	High Yield US	0.39	70%*	53%*	66%*	78%*
Europe	Eurostoxx	0.79	92%*	90%*	90%*	95%*
	MSCI Small Cap	0.55	83%*	67%*	79%*	92%*
	Investment Grade EMU	0.10	33%*	29%*	51%*	33%*
	High Yield EMU	0.27	61%*	49%*	63%*	70%*
Emerging	MSCI EM Europe	0.30	65%*	50%*	57%*	75%*
	MSCI EM Asia	0.33	65%*	51%*	64%*	77%*
	MSCI EM Latam	0.46	77%*	63%*	79%*	87%*
Currencies	USDCHF	0.00	-2%	28%*	-3%	-20%*
	EURUSD	0.01	19%*	-18%*	19%*	44%*
	USDJPY	0.03	25%*	13%*	10%	39%*
	GBPUSD	0.00	20%*	-19%*	15%*	42%*
	USDNOK	0.03	-31%*	14%*	-19%*	-60%*
	USDBRL	0.19	-50%*	-25%*	-49%*	-74%*
	USDSGD	0.06	-37%*	3%	-30%*	-59%*
Commodities	DJ Commodity	0.07	38%*	8%	18%*	59%*
	GSCI Agri.	0.03	26%*	1%	20%*	39%*
	GSCI Energy	0.04	27%*	0%	-6%	57%*
	GSCI Ind. Metals	0.11	49%*	31%*	34%*	64%*
	GSCI Prec. Metals	0.00	11%*	-5%	23%*	14%*
	Gold	0.00	6%	-6%	22%*	7%
Rates	US 2Y Rate	0.15	40%*	49%*	36%*	37%*
	US 10Y Rate	0.22	49%*	49%*	44%*	50%*
	German 2Y Rate	0.18	39%*	45%*	40%*	40%*
	German 10Y Rate	0.20	49%*	39%*	45%*	54%*
	Inv. Grade US OAS	0.18	-47%*	-31%*	-52%*	-57%*
	High Yield US OAS	0.30	-64%*	-51%*	-63%*	-76%*
	Inv. Grade EMU OAS	0.16	-41%*	-35%*	-42%*	-49%*
	High Yield EMU OAS	0.21	-55%*	-45%*	-56%*	-65%*
Vol. Indices	VIX	0.42	-71%*	-67%*	-77%*	-75%*
	VDAX	0.42	-68%*	-66%*	-75%*	-69%*

Table 2: Correlation analysis of the returns considered in the dataset.

The statistics presented in the table are computed using logarithmic returns over the period that starts on 08/01/1998 and ends on 12/31/2011. “*” indicates a correlation coefficient statistically different from zero à the 5% level. R2 refers to the R^2 of the regression of the returns on the first factor of the PCA.

A.2 Empirical Results

	Model Averaging S1 Mean	Model Averaging S2 Mean	Model Averaging S1 Vol.	Model Averaging S2 Vol.
SP500	-2.211	-22.287	16.953	38.926
Russell 2000	4.683	-31.17	21.67	46.097
Investment Grade US	1.151	-3.059	1.678	7.66
High Yield US	3.486	-11.048	5.844	20.284
Eurostoxx	-1.918	-35.4	19.444	44.225
MSCI Small Cap	8.053	-48.99	17.883	43.923
Investment Grade EMU	0.748	-5.373	1.602	5.501
High Yield EMU	2.589	-16.125	8.408	22.511

Table 3: Asset annual average return (in %) and volatility for each state.

Threshold	1999-2012		1999-2004		2005-2007		2008-2011	
	AG Test	IS Returns	AG Test	IS Returns	AG Test	IS Returns	AG Test	IS Returns
0.1		1.365		0.313		-0.449		2.768
0.2	8.429*	1.098	7.708*	0.313	7.928*	-0.449	1.647*	1.895
0.3	5.851*	1.154	6.345*	0.394	6.164*	-0.449	0.259*	1.363
0.4	5.211*	1.785	5.797*	0.521	5.404*	2.137	0.221*	0.332
0.5		2.955		0.222		2.137		6.864
0.6	-1.919	2.224	-3.139	-0.229	-5.793	3.732	1.22*	7.378*
0.7	-1.015*	2.927	-1.602*	0.551	-5.802	5.224	1.58*	0.245
0.8	-0.29*	5.39*	-0.199*	0.771	-5.963	9.891*	1.82*	2.767
0.9	0.567*	4.126	0.964*	1.005*	-5.413	4.803*	2.145*	0.346

Table 4: In-sample econometric and economic threshold effect.

This table displays the Vuong's in sample density test comparing the likelihood of the model considering that the financial markets are optimist when π_t is greater than 0.5 and the likelihood of the models taking into account various threshold, from 0.2 to 0.9. A "*" indicates that 0.5 provides the best or an equivalent fit than the competitor. This table also provides the annualized return obtained for various threshold indicator. For instance, a threshold equal to 0.6 indicates that the market portfolio is bought if π_t is above 0.6 and sold otherwise.

		Sub-sample	Horizon	Against 0.3	Against 0.4	Against 0.6	Against 0.7	Against 0.8
AG Test	2008-2011	T+1	2.199	1.183	1.451	2.382	2.807	
		T+2	1.416	1.408	0.796	2.042	2.628	
		T+3	2.913	1.824	0.622	1.949	2.477	
		T+4	1.484	1.702	1.076	2.217	2.721	
	2008-2009 2010-2011	T+1	2.15	1.238	1.575	2.362	2.693	
		T+1	0.66	-0.321	-0.428	0.44	0.908	
		OS Return Comparison	T+1	1.982	1.263	1.982	1.728	0.275
			T+2	5.197	2.963	5.197	-1.069	0.451
T+3	6.903		5.941	6.903	1.385	3.005		
T+4	10.836		8.285	10.836	0.57	-1.424		
2008-2009 2010-2011	T+1	2.508	0.399	2.508	-0.42	-1.195		
	T+1	3.607	2.107	3.607	3.849	1.767		

Table 5: Out-of-sample econometric and economic threshold effect.

The top part of this table displays the Amisano and Giacomini's (2007) out-of-sample density test comparing the likelihood of the model considering that financial markets are optimist when π_t is greater than 0.5 and the likelihood of models taking into account various thresholds, from 0.2 to 0.8. This table also provides the difference in the annualized returns between a strategy driven by the 0.5 threshold and the strategies driven by various ones, from 0.2 to 0.8.

Weight function	Normal 1999-2011	Right Tail 1999-2011	Left Tail 1999-2011	Normal 1999-2004	Normal 2005-2007	Normal 2008-2011
Against RAI-KP	5.472	4.119	5.058	4.976	6.327	2.494
Against RAI-MI	4.398	4.662	4.052	6.463	7.624	1.954
Against Multivariate Approach	-4.479	-5.08	-4.188	-2.355	-10.832	-0.836
Against PCA	2.591	2.93	2.396	0.952	5.499	1.753

Table 6: Vuong's in sample density test.

		Buy and Hold	Model Averaging	RAI-KP	RAI-MI	PCA Approach	Multivariate Approach
Returns by periods	1999-2012	-1.056	2.955*	0.322	-0.823	-0.066	1.657
	1999-2004	0.313	0.222	-0.51	-0.557	0.667*	0.096
	2005-2007	-0.449	2.137	4.188*	-1.877	-0.449	4.093
	2008-2011	-4.959	6.864*	1.995	4.199	-4.959	2.103
Descript. Statistics	Vol.	10.096	10.096	10.096	10.096	10.096	10.096
	Skewness	-0.755	0.18	0.227	0.345	-0.667	0.274
	Kurtosis	2.579	2.101	2.126	2.107	2.624	2.099

Table 7: In sample annualized returns (in %) obtained from the different measures of probability.

<i>Weight function</i>	<i>Sub-sample</i>	<i>Horizon</i>	<i>Against RAI-KP</i>	<i>Against RAI-MI</i>	<i>Against Multi.</i>	<i>Against PCA</i>
Normal	2008-2011	T+1	3.023*	3.168*	2.076*	2.673*
		T+2	3.349*	3.198*	3.186*	2.683*
		T+3	3.927*	3.566*	4.281*	3.236*
		T+4	3.67*	3.634*	4.787*	2.788*
	2008-2009 2010-2011	T+1	2.747*	2.866*	2.764*	2.668*
Right Tail	2008-2011	T+1	1.607	1.435	-1.029	0.907
		T+2	2.081*	2.249*	2.003*	0.97
		T+3	2.383*	2.402*	2.535*	1.553
		T+4	1.887	1.956	2.084*	1.322
	2008-2009 2010-2011	T+1	1.43	1.224	2.287*	1.073
Left Tail	2008-2011	T+1	2.901*	3.051*	2.169*	2.612*
		T+2	3.154*	3.017*	3.015*	2.632*
		T+3	3.901*	3.559*	4.328*	3.29*
		T+4	3.546*	3.517*	4.706*	2.745*
	2008-2009 2010-2011	T+1	2.634*	2.763*	2.656*	2.578*
		T+1	1.912	1.867	-1.985	0.583

Table 8: Amisano and Giacomini's (2007) out-of-sample density test.

Horizon	Model	2008-2011	2008-2009	2010-2011
T+1	Buy and Hold	-3.991	-5.984	-2.811
	Model Averaging	5.649	4.298	7.923
	RAI-KP	-0.719	-0.824	-1.624
	RAI-MI	3.827	4.856	1.568
	Multivariate M.	4.069	3.752	5.337
	PCA	-3.621	-5.273	-2.811
T+2	Buy and Hold	-3.745	-4.522	-2.527
	Model Averaging	3.978	6.68	0.888
	RAI-KP	3.175	4.574	1.947
	RAI-MI	-1.122	1.648	-3.5
	Multivariate M.	-0.331	1.075	-1.983
	PCA	-0.874	1.128	-2.527
T+3	Buy and Hold	-3.724	-3.834	-2.831
	Model Averaging	6.165	9.443	2.065
	RAI-KP	2.871	3.353	2.932
	RAI-MI	5.749	12.049	0.205
	Multivariate M.	5.95	5.146	5.822
	PCA	-1.213	1.133	-2.831
T+4	Buy and Hold	-3.375	-4.355	-1.919
	Model Averaging	1.653	6.4	-3.273
	RAI-KP	-0.461	0.386	-0.958
	RAI-MI	-13.673	-18.219	-7.994
	Multivariate M.	2.794	5.897	-0.665
	PCA	-3.765	-5.108	-1.919

Table 9: Out-of-sample annualized returns obtained from the different measures of probability.

A.3 Extension - Additional assets global comparison

	In-Sample Analysis		Out-of-Sample Analysis	
	Density Test	Economic Difference	Density Test	Economic Difference
Emerging	2.742*	0.762	1.816	0.847
MSCI EM Asia	-0.455	-0.735	0.666	-1.924
MSCI EM Latam	1.632	-0.041	0.91	-0.54
MSCI EM Europe	2.193*	0.241	1.28	-1.704
Currencies	4.348*	0.038	1.08	1.994
USDCHF	1.699	-0.041	0.447	-1.248
EURUSD	-1.391	0.036	1.078	0.719
USDJPY	7.967*	0.074	1.961*	0
GBPUSD	1.231	-0.102	0.727	-1.319
USDNOK	6.431*	0.074	1.141	1.063
USDBRL	2.018*	0.249	-1.316	-1.53
USDSGD	1.632	-0.07	1.449	-1.409
Commodities	6.208*	0.38	2.355*	2.468
DJ Commodity	1.72	-0.029	-0.811	-1.103
GSCI Agri.	1.771	0.507	1.319	-0.165
GSCI Energy	1.959	0.082	0.698	0.41
GSCI Ind. Metals	4.136*	0.497	0.407	1.715
GSCI Prec. Metals	2.704*	-0.51	0.543	1.535
Gold	3.16*	-0.117	1.411	0.645
Rates	7.874*	-0.253	0.415	0.863
US 2Y Rate	2.007*	0.143	-0.35	-0.644
US 10Y Rate	4.895*	-0.443	0.506	0.623
German 2Y Rate	4.539*	0.569	0	-0.489
German 10Y Rate	4.18*	-0.243	0.946	-0.077
Inv. Grade US OAS	2.946*	-0.457	-1.25	-1.263
High Yield US OAS	0.714	0.003	-0.467	-0.096
Inv. Grade EMU OAS	2.165*	-0.31	-1.227	-0.077
High Yield EMU OAS	-1.289	-0.73	0.192	-1.902
Vol	3.199*	0.076	-0.188	-0.218
VIX	4.865*	0.164	0.808	-0.795
VDAX	1.924	-0.025	-0.963	-0.008

Table 10: In- and Out-of-Sample analysis – both econometric and economic approaches – of the introduction of additional assets to the core sample.

This table displays the consequences of an additional asset in the core sample. Only the results relative to the model averaging approach are presented. For example, the density test value of 1.699 for the USDCHF corresponds to the Vuong's statistics value of the density test comparing the loglikelihood computed with the core sample and the one obtained with the core sample and the Swiss Franc against US Dollar. Similarly, an economic difference of -0.041 means that the sample built with the core assets and the Swiss Franc against the US Dollar allows to obtain 0.041% of additional yearly returns compared to the core sample. For *Emerging*, all the emerging assets – MSCI EM Asia, MSCI EM Latam and MSCI EM Europe – are added to the core sample. A similar approach is adopted for the currencies, the commodities, the rates and the volatility indices.

	Normal 1999-2011	Right Tail 1999-2011	Left Tail 1999-2011	Normal 1999-2004	Normal 2005-2007	Normal 2008-2011
Emerging	2.742*	4.266*	2.115*	2.223*	6.005*	0.061
Currencies	4.348*	5.04*	3.871*	3.637*	6.033*	0.318
Commodities	6.208*	6.644*	5.895*	6.15*	6.267*	0.809
Rates	7.874*	7.2*	7.185*	4.524*	31.692*	2.976*
Vol. Indices	3.199*	5.036*	2.465*	1.000	6.989*	1.355

Table 11: Vuong's in sample density test – Core assets against alternative samples.

		Core Data	Emerging	Currencies	Commodities	Rates	Vol. Indices
Returns by periods	1999-2012	6.674	6.582	6.251	6.044	7.447*	6.722
	1999-2004	0.634	0.898	0.884	0.851	1.048*	0.981
	2005-2007	10.911	10.911	8.883	10.911	11.589*	10.911
	2008-2011	8.598*	5.837	6.372	4.545	6.747	5.529
Descript. Statistics	Vol.	10.096	10.096	10.096	10.096	10.096	10.096
	Skewness	0.569	0.591	0.591	0.608	0.578	0.604
	Kurtosis	6.506	6.501	6.498	6.494	6.513	6.497

Table 12: In sample annualized returns obtained from the different samples.

<i>Weight function</i>	<i>Sub-sample</i>	<i>Horizon</i>	<i>Emerging</i>	<i>Currencies</i>	<i>Commodities</i>	<i>Rates</i>	<i>Vol. Indices</i>
Normal	2008-2011	T+1	1.816	1.080	2.355*	0.415	-0.188
		T+2	1.272	0.741	2.177*	-0.468	-0.873
		T+3	2.253*	2.143*	2.127*	-0.158	0.274
		T+4	2.344*	1.362	2.358*	-0.557	-0.22
	2008-2009	T+1	1.764	0.925	2.24*	0.188	-0.553
	2010-2011	T+1	0.828	1.061	1.248	1.263	1.984*
Right Tail	2008-2011	T+1	1.797	2.084*	0.869	1.732	2.956*
		T+2	1.326	1.541	1.226	2.014*	2.202*
		T+3	1.666	1.841	1.508	0.347	1.345
		T+4	1.056	-0.669	1.319	0.624	-0.083
	2008-2009	T+1	1.695	1.744	0.269	1.528	2.067*
	2010-2011	T+1	1.100	2.344*	3.705*	1.727	2.534*
Left Tail	2008-2011	T+1	1.56	0.774	2.314*	-0.097	-0.472
		T+2	1.105	0.504	2.143*	-0.706	-1.157
		T+3	2.068*	2.056*	2.112*	-0.235	0.104
		T+4	2.652*	1.434	2.345*	-0.747	-0.219
	2008-2009	T+1	1.524	0.669	2.253*	-0.288	-0.746
	2010-2011	T+1	0.56	0.741	0.824	1.01	1.527

Table 13: Amisano and Giacomini's (2007) out-of-sample density test – Core Assets against alternative samples.

<i>Horizon</i>	<i>Model</i>	<i>2008-2011</i>	<i>2008-2009</i>	<i>2010-2011</i>	<i>Vol.</i>	<i>Skewness</i>	<i>Kurtosis</i>
T+1	Core Dat	5.649	4.298*	7.923	10.096	0.711	2.361
	Emerging	4.802	3.585	6.989	10.096	0.762	3.356
	Currencies	3.655	2.333	3.719	10.096	0.713	3.401
	Commodities	3.181	0.099	7.39	10.096	0.764	3.4
	Rates	4.786	2.969	7.598	10.096	0.775	3.355
	Vol. Indices	5.867*	3.703	9.004*	10.096	0.731	3.339
T+2	Core Dat	3.978	6.68	0.888	10.096	0.46	2.518
	Emerging	4.293	6.768	1.356	10.096	0.561	3.421
	Currencies	4.445	7.525*	1.597	10.096	0.56	3.419
	Commodities	1.41	2.129	0.306	10.096	0.605	3.477
	Rates	3.348	4.596	1.636	10.096	0.561	3.44
	Vol. Indices	4.619*	6.061	2.659*	10.096	0.554	3.417
T+3	Core Dat	6.165	9.443	2.065	10.096	0.414	2.513
	Emerging	7.745*	12.07	2.579	10.096	0.526	3.405
	Currencies	7.145	12.322*	2.579	10.096	0.54	3.403
	Commodities	1.531	0.492	1.848	10.096	0.388	3.526
	Rates	4.517	7.651	0.619	10.096	0.613	3.415
	Vol. Indices	7.142	10.174	3.206*	10.096	0.552	3.4
T+4	Core Dat	1.653	6.4	-3.273	10.096	0.481	2.524
	Emerging	4.968*	10.136*	-0.666	10.096	0.536	3.385
	Currencies	2.735	7.502	-1.444	10.096	0.595	3.414
	Commodities	-2.104	-4.877	0.343*	10.096	-0.686	4.068
	Rates	4.631	9.084	-0.338	10.096	0.568	3.384
	Vol. Indices	3.745	9.011	-1.917	10.096	0.59	3.395

Table 14: Out-of-Sample sample annualized returns obtained from the different samples.

A.4 Extension - Macroeconomic data global comparison

	In-Sample Analysis		Out-of-Sample Analysis	
	Density Test	Economic Difference	Density Test	Economic Difference
Business Climate	2.666*	-0.716	1.134	2.228
ISM US	3.788*	-0.993	-0.246	2.228
IFO Germany	1.715	-0.272	0.516	2.33
Industry Survey EMU	0.18	-0.009	1.645	2.228
Unemployment	1.208	0.609	0.609	2.228
Non Farm Payrolls	-0.42	-0.44	1.569	2.228
Unemployment Rate US	0.078	-0.602	0.728	1.45
Unemployment Rate EMU	2.211*	0.668	0.039	1.45
CPI	2.356*	-1.283	1.083	-2.991
CPI US	1.702	0.201	-0.798	2.33
CPI EMU	0.809	-0.09	1.536	0.017
Total	3.042*	-0.895	2.583*	-4.966

Table 15: In- and Out-of-Sample analysis – both econometric and economic approaches – of the introduction of macroeconomic data to the core sample.

This table displays the consequences of an additional macroeconomic data in the core sample. Only the results relative to the model averaging approach are presented. For example, the density test value of 3.788 for the ISM US corresponds to the Vuong's statistics value of the density test comparing the loglikelihood computed with the core sample and the one obtained with the core sample and the ISM US. Similarly, an economic difference of -0.993 means that the sample built with the core assets and ISM US allows to obtain 0.993% of additional yearly returns compared to the core sample. For *Business Climate*, all the macroeconomic data relative to the business climate – ISM US, IFO Germany and Industry Survey EMU – are added to the core sample. A similar approach is adopted for the Unemployment and the CPI. The Total is obtained by considering all the macroeconomic data.

B Figures

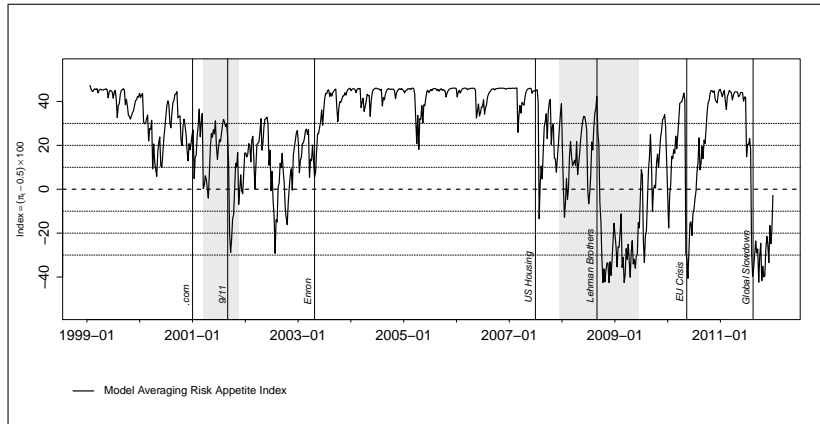


Figure 1: Risk Appetite Index from 1998 to 2011.

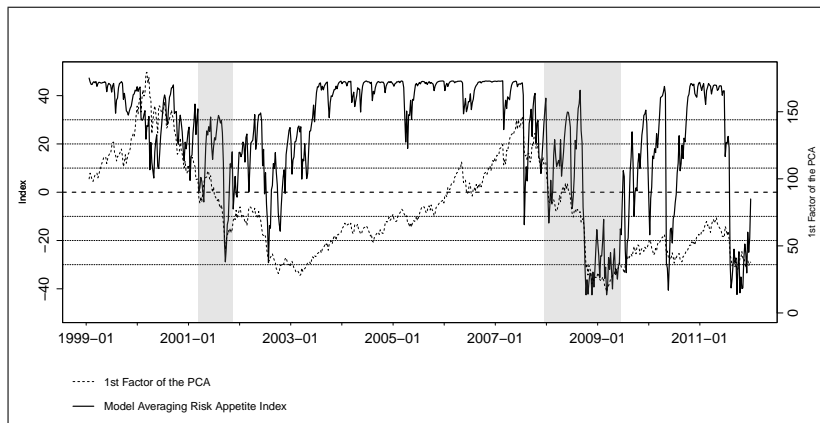


Figure 2: Risk Appetite Index and 1st Factor of the PCA from 1998 to 2012.

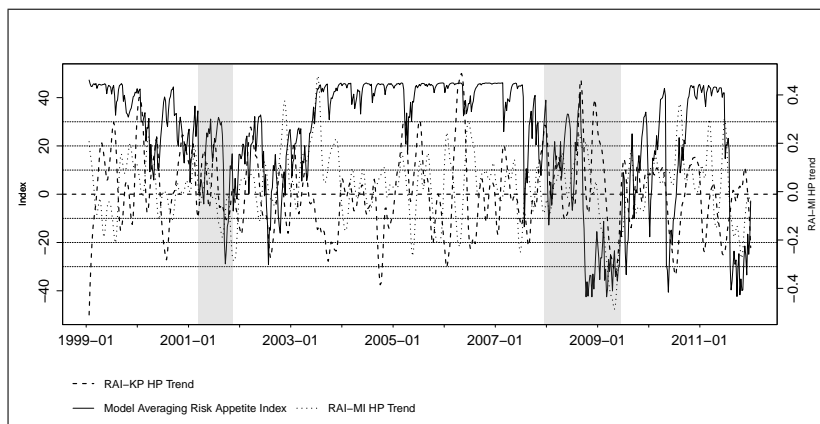


Figure 3: Model Averaging, Kumar-Persaud and Misina risk appetite approaches.

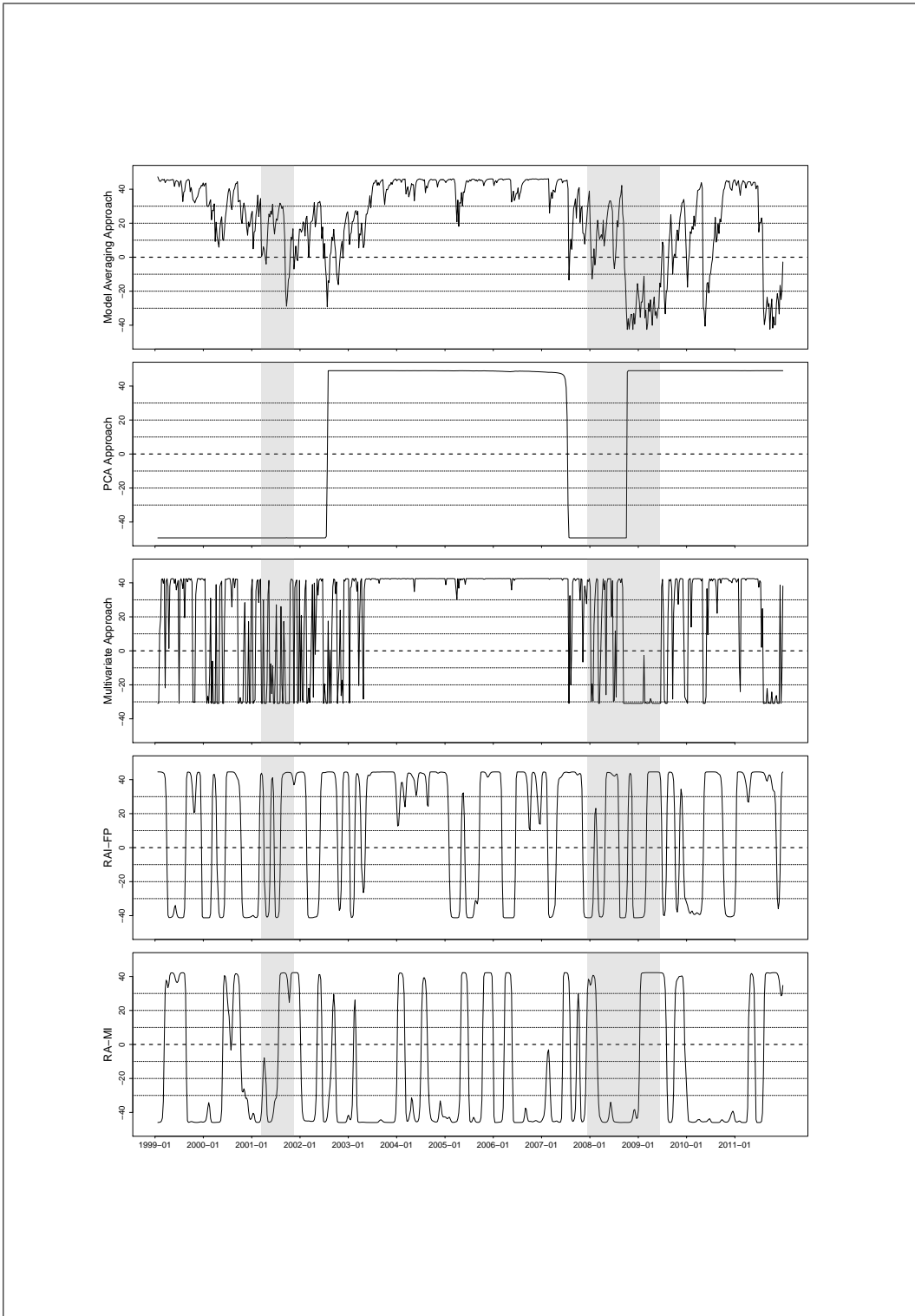


Figure 4: Model Averaging, First Factor and Multivariate Risk Appetite Indices from 1998 to 2011.

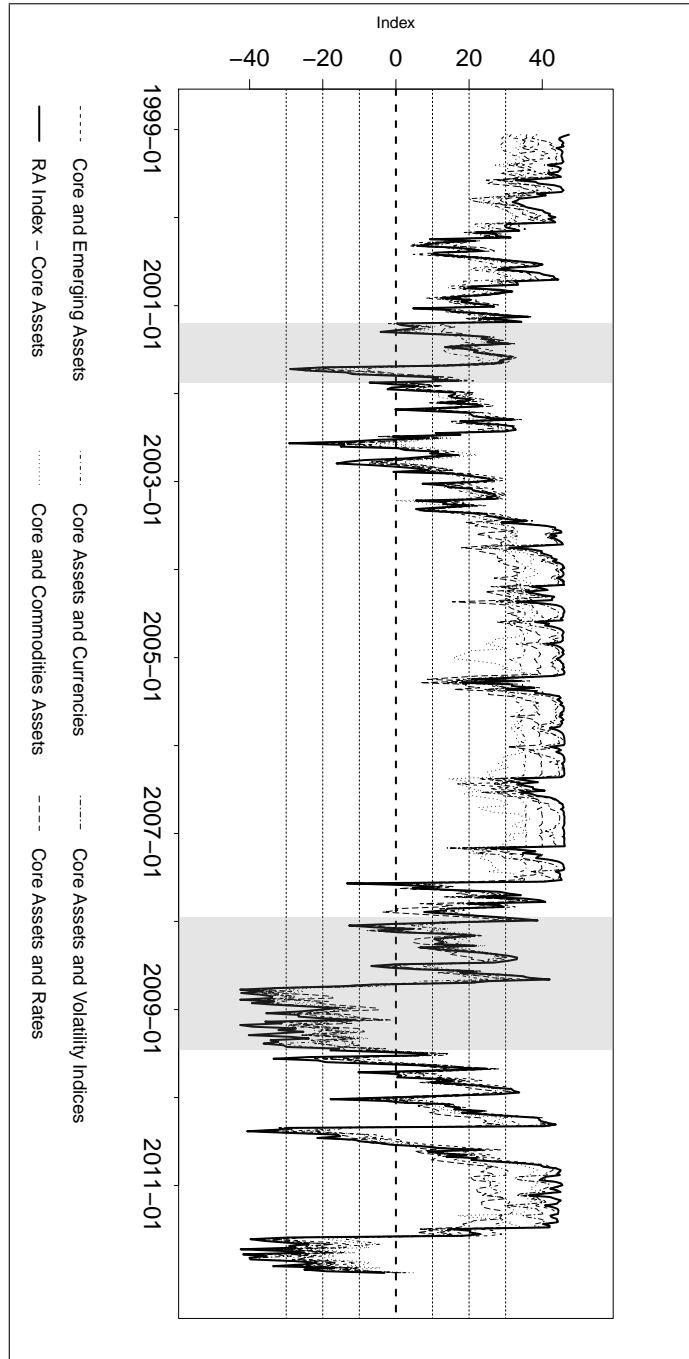


Figure 5: Risk Appetite Evolution for Core Assets and Alternative Samples from 1998 to 2011.

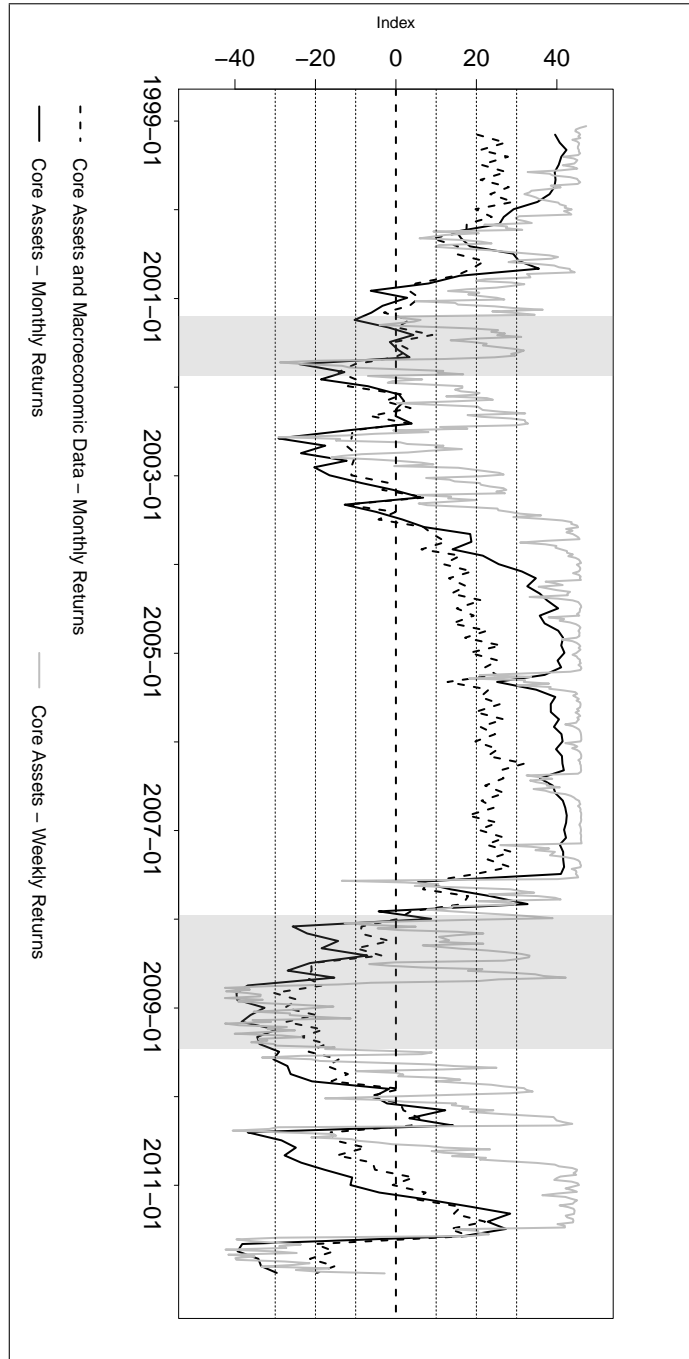


Figure 6: Risk Appetite Evolution for Core Assets and Macroeconomic Data from 1998 to 2011.