BAFFI CAREFIN
Centre for Applied Research on International Markets, Banking, Finance and Regulation

Working Paper

Marta Giampietro, Massimo Guidolin, Manuela Pedio

Can No-Arbitrage SDF Models with Regime Shifts Explain the Correlations Between Commodity, Stock, and Bond Returns?

BAFFI CAREFIN Centre Research Paper Series
No. 2016- 19

Can No-Arbitrage SDF Models with Regime Shifts Explain the Correlations Between Commodity, Stock, and Bond Returns?[♣]

Marta Giampietro
Royal Bank of Canada

Massimo Guidolin
Bocconi University, BAFFI CAREFIN Centre

Manuela Pedio
Bocconi University

December 2015

Abstract

We investigate whether it is possible to find a Stochastic Discount Factor (SDF) that jointly prices the cross-section of eight U.S. portfolios of stocks, Treasuries, corporate bonds, and commodities and replicates their observed moments, and especially correlations. We use the first three principal components extracted from a set of 112 U.S. macro variables as pricing factors. We also introduce commodity-based factors in the SDF, motivated by commodity pricing theories and we use them either separately or in conjunction with the macro-based ones. We report that it is not impossible to find a set of macroeconomic factors able to jointly price the cross section of stocks, bonds, and commodities; however, this task is accomplished by a small set of commodity-based factors. Observed correlations for commodities are perfectly matched by a parsimonious, single state, diagonal factor VAR model where only three commodity-based factors enter the SDF. While introducing regimes does not improve the performance in all cases, they could be beneficial to replicating correlations among commodity and bond portfolios.

[♣] Correspondence to Massimo Guidolin, massimo.guidolin@unibocconi.it.

1. Introduction

The interest of institutional investors in commodities has increased significantly over the past few years. In addition, in recent years, thanks to the role played by exchange traded funds, also retail investors have developed a growing taste for this asset class. This is due to the fact that commodities can generate equity-like returns in the long-run, act as risk diversifiers both in the short- and in the long-run, and serve as an inflation hedge (see, e.g., Erb and Harvey, 2006; Gorton and Rouwenhorst, 2006).[1] In spite of their growing importance in institutional and retail portfolios, our understanding of this asset class remains unsatisfactory. While the asset pricing literature is extensive as far as traditional asset classes, especially equities, are concerned, there is no consensus on whether there is an asset pricing model which can effectively explain both the cross-sectional and the time series variation in commodity returns (see, e.g., the discussion in Bakshi, Gao and Rossi, 2014). Part of the reason of this lack of in-depth understanding of commodities, is that they require the development of new methods and operational approaches.

Our paper explores one such combination of novel methods—Hidden Markov chain models (HMM) to describe time heterogeneity (see Dias et al., 2015) in the fundamental pricing measure governing the cross-section of asset returns—applied to a number of questions concerning commodities as an asset class. We also develop a methodology to operationalize the estimation of HMM stochastic discount factors (henceforth, HMMSDF) that relies on variations of standard Expectations-Maximization algorithms to maximize the log-likelihood function.[2]

When applied to commodities, the classical asset pricing research issues—for instance the nature of the factors driving their risk premia and volatility—have been approached in a number of ways: some researchers have employed commodity-specific factors (see, e.g., Szymanowska, de Roon, Nijman and Goorbergh, 2013; Yang, 2013; Bakshi, et al., 2014, among others); other authors have used models designed to price all assets (see, e.g., Asness, Moskowitz and Pedersen, 2013; Koijen, Pedersen, Moskowitz and Vrugt, 2013, among others). The former approach considers commodities as a separate asset class for which specific pricing factors are needed; the latter assumes that financial markets are perfectly integrated and hence that a unique pricing measure that prices all assets can be found. Moreover, most studies have tested asset pricing models on commodity return data in a stand-alone fashion while few others have augmented the test asset menu with other asset classes (see e.g., Dhume, 2011; Asness et al., 2013).

Amid this heterogeneity of approaches, a conclusion as to the best asset pricing model for commodities has not been found yet. There are (at least) two important issues

[1] Recent research has shown that momentum and term structure strategies work well with commodities, suggesting that they should also become part of the tactical asset allocation of investors (see, e.g., Miffre and Rallis, 2007).

[2] SDF models are based on the result that in the absence of arbitrage, the price of an asset at time t is the expected discounted value of the asset payoff in period $t+s$, based on information available at time t, when discounting and integration are performed under the risk neutral measure implied by the SDF.

which remain unresolved: first, whether commodities form an alternative asset class (see Gorton and Rouwenhorst, 2006) and, as such, they may be segmented from other financial assets; second, whether commodity returns are characterized by a heterogeneous structure (as in Erb and Harvey, 2006 and Kat and Oomen, 2007) in which case, it would be impossible to find a set of common factors able to price their entire cross-section. Commodities display in fact many differences: some are storable and some are not, some are input goods and some are intermediate goods.

To explore these questions, we develop methods useful to investigate whether it is possible to find a SDF that jointly prices the cross-section of stock, bond, and (spot) commodity returns and that replicates the empirically observed moments of returns, with a specific interest for the matrix of correlations among pairs of commodities as well as commodities and other asset classes. Our candidate SDFs are linked to the state of the economy through a rich set of variables concerning prices, production, and the labor and housing markets. We employ standard linear projections of latent SDFs on principal components that summarize such variables. Our set of tested models are then further extended to include latent regime shifts governed by one ergodic and irreducible Markov chain driving shifts in the coefficients that map the K priced factors into the SDF.

In an application to commodities research, we use a cross section of eight portfolios of stocks, Treasuries, corporate bonds, and commodity indices (the S&P-GS Agriculture and Livestock, Precious Metals, Industrial Metals, and Energy indices.) over the period January 1989 – December 2011 as test assets to estimate parametric SDFs designed to jointly price all portfolios. Besides using the principal components extracted from a large set of macroeconomic variables (see Ludvigson and Ng, 2009), we also augment the set of factors by introducing commodity-specific factors motivated by commodity pricing theories and we use them either separately or in conjunction with the macro-based factors. Such an empirical exercise allows us to answer the two questions stated above. The possibility (or impossibility) of finding a unique SDF that jointly prices the cross section of stock, bond, and commodity returns would indicate that commodity markets are integrated with (segmented from) the other markets. The possibility (or impossibility) of finding a SDF that prices the cross-section of commodity returns only would point to a homogeneous (heterogeneous) pricing structure for commodities.

To the best of our knowledge, also our applied contribution is original for (at least) two reasons. First, when we extend our model to include latent regime shifts governed by one ergodic and irreducible Markov chain, we test whether there is evidence of a need of additional parametric flexibility to price a range of asset classes that includes commodities. It is generally accepted that financial markets follow boom-and-bust cycles that involve both the mean and the volatility of asset returns (see Guidolin, 2011 and 2012). Recently, the awareness that also correlations between returns on different asset classes would undergo massive changes has emerged (see Ang and Bekaert, 2002; Bae et al., 2014). While the literature has mainly focused on stocks and bonds so far (see, e.g., Guidolin and Timmermann, 2006), the number of researchers who has applied Hidden Markov models (HMM) to commodities is limited (see, e.g., Alizadeh, Nomikos, Poulialis, 2014; Bae et al., 2014; Lee and Yoder, 2007*).* We contribute to this strand of research even though our focus is not simply on the modeling of regimes and time-varying moments in commodity returns, but the on time variation that is induced by the presence of regime shifts in the relationships between the SDF and an interpretable set of priced risk factors.

The second contribution of our work is that we look for an SDF that not only prices the cross-section of stock, bond, and commodity returns in terms—as it is customary—of their means and variance (hence, Sharpe ratios), but that also replicates their observed correlations. This is particularly useful in the light of the recent developments in the commodity markets. While there is some evidence that prior to the early 2000s, commodities shared co-movements with stocks (see Gorton and Rouwenhorst, 2006) or each other (see Erb and Harvey, 2006), it seems well established in the minds of investors and researchers that the commodity markets have recently undergone deep changes: commodity prices have experienced booms followed by significant busts and the correlation between stock and commodity returns has been shown to have risen dramatically since mid-2008 (see, e.g., Miffre, 2011; Tang and Xiong, 2012; Büyüksahin and Robe, 2013).[3]

Our empirical results can be summarized as follows. First, SDFs based on commodity-related factors outperform all other SDFs both in terms of statistical fit and for their ability to match the sample moments that characterize commodity data, in particular pair-wise correlations. Therefore our paper proposes afresh the classical puzzle that factors derived from macroeconomic variables fail to fit the most important stylized facts concerning asset returns, including commodities. However, some modest pay-off in terms of the ability of models to fit the data is obtained when both macro- and commodity-based factors are used jointly. This finding implies that commodity markets are to some extent segmented from the remaining asset markets investigated, because specific factors are required to obtain an effective pricing performance. Second, even though on statistical grounds there is some evidence that regimes may characterize the best performing SDFs, we find that the ability of HMM pricing measures to improve over the corresponding single-state SDFs is limited, although a full understanding of the (dynamic) correlations between bond portfolios and commodities requires that regimes are taken into account.

[3] There is no consensus as to the explanations for this evidence but many commentators agree on the financialization of commodities as a possible cause (see, e.g., Basak and Pavlova, 2013; Büyükşahin and Robe, 2013; Singleton, 2012). According to this hypothesis, the increase in investments in commodity-related products, especially commodity indices, has caused these strong swings in commodity prices. In fact, such an increase in correlations may be sufficiently strong that the diversification benefits of commodities may be put into question (see, e.g., Daskalaki and Skiadopoulos, 2011).

Two papers are closely related to ours. Bakshi, et al. (2014) show that three commodity factors, namely an average commodity factor, a carry factor, and a momentum factor are able to predict output growth, bond returns, equity returns, and returns on commodity currencies. Specifically, their predictive regressions support the view that the commodity factors can forecast real GDP growth; they also report that the factors can forecast bond and equity returns up to 12 months, with predictive slope coefficients that are mutually compatible across GDP growth, bond, and equity returns. Similarly to their paper, we use commodity factors to specify SDFs used to price a wide range of assets and that therefore must reflect underlying business cycle conditions. However, while their focus revolves on the direct predictive performance for growth and asset returns, in our paper the factors affect instead the SDF. Basu and Miffre (2013) construct factor mimicking portfolios to examine the systematic effects of hedging pressure on the commodity futures risk premium. Using single- and double-sorts they systematically buy the contracts for which hedgers are the shortest and/or speculators are the longest and sell the contracts for which hedgers are the longest and/or speculators are the shortest. They find empirical support to the hypothesis that hedging pressure is a systematic factor. They also use the two-step Fama and MacBeth's method to investigate whether the hedging pressure, momentum and term structure portfolios explain cross-sectional commodity futures returns. While momentum is not priced cross-sectionally, the prices of risk associated with hedging pressure and the term structure are found to be significantly positive. However, they only focus on pricing commodities, while we formalize and estimate an SDF that depends on two of the three factors in Basu and Miffre's work.

The rest of the paper has the following structure. Section 2 introduces SDF models. Linear factor SDF models are presented in section 2.1 and are generalized to the HMM case in section 2.2. Section 2.3 presents the maximum likelihood estimation strategy and explains in detail how the Baum-Welch algorithm can be used to implement it. Section 3 introduces our application. Section 4 contains our key empirical findings: first we conduct an extensive search to find the best model specification and show the estimates of the best fitting models; next, we assess their comparative ability to match empirical correlations. Section 5 concludes.

2. The Methodology: Factor-Based SDF Models

SDF models provide a general framework for pricing assets: many existing asset pricing methods, such as the capital asset pricing model, the general equilibrium, consumption-based, inter-temporal capital asset pricing model, and also Black and Scholes' formula, can all be shown to be specializations of SDF models to reflect specific assumptions. One SDF is defined by a simple proposition: the price of an asset at time *t* is equal to the expected payoff of the asset in t+1, based on information available at time t:

$$p_{i,t} = E\left[M_{t+1}X_{i,t+1}|\Im_t\right] \qquad i = 1, \ldots, n, \qquad (1)$$

where $p_{i,t}$ is the price of the ith asset at time t, $X_{i,t+1}$ is the asset payoff in $t+1$, M_{t+1} is the SDF (a random variable), and $E[\cdot\,|\Im_t]$ indicates the expectation based on the information available at time t. The existence of an SDF is equivalent to the law of one price, its positivity is equivalent to the absence of arbitrage opportunities, and its uniqueness is equivalent to market completeness.

If we divide both sides of equation (1) by $p_{i,t}$, we obtain

$$1 = E\left[M_{t+1}\frac{X_{i,t+1}}{p_{i,t+1}}|\Im_t\right] = E\left[M_{t+1}\left(1 + R_{i,t+1}\right)|\Im_t\right], \qquad (2)$$

where $1 + R_{i,t+1} = X_{i,t+1}/p_{i,t+1}$ is the gross return of asset i. (2) is the standard Euler condition under a generic SDF, M_{t+1}, derived in a representative agent framework. We define $m_{t+1} \equiv \ln M_{t+1}$, $r_{i,t+1} \equiv \ln(1+R_{i,t+1})$ and $u_{i,t+1} \equiv m_{t+1} + r_{i,t+1}$. If the random vector $\boldsymbol{y}_{t+1} \equiv [m_{t+1}\ r_{1,t+1}\ r_{2,t+1} \ldots r_{n,t+1}]'$, with $\Psi_t \equiv \{\boldsymbol{y}_{t-s}\}_{s\geq 0} \subset \Im_t$, has a stationary, homoskedastic multivariate Gaussian distribution, Appendix A shows that

$$E\left[r_{i,t+1}|\Psi_t\right] = -E[m_{t+1}|\Psi_t] - \frac{1}{2}Var[m_{t+1}] - \frac{1}{2}Var\left[r_{i,t+1}\right] - Cov\left[r_{i,t+1}, m_{t+1}\right], \qquad (3)$$

which establishes a functional link between conditional risk premia on any asset or portfolio, the corresponding first and second moments of the assumed SDF, and the covariance between such an operator and the net returns on the very asset. Moreover, this general expression can be specialized further when assumptions are imposed on the SDF functional form. In the following, we distinguish between two cases, the single- vs. the multi-state, HMM case. In the following, we do not take a stance as to whether the SDF is unique or not, but we focus—under any assumed model structure—to the SDF that maximizes some statistical criterion and therefore emerges from the data. Therefore, in the case of market incompleteness (as in Marroquìn-Martinez and Moreno, 2013), we assume that the portfolios to be priced are replicable, so that their equilibrium return does not depend on the SDF selected.

2.1. Linear Factor SDF Models

Suppose the SDF has the log-linear structure,

$$M_{t+1} = exp\left(\gamma_0 + \sum\nolimits_{j=1}^{K}\gamma_j f_{j,t+1}\right) > 0, \qquad (4)$$

where $f_{1,t+1}, f_{2,t+1}, \ldots, f_{K,t+1}$ are K systematic factors driving the way in which all assets are priced. In the following, also for identification purposes when the model is generalized to the HMM case, we assume that the K factors are observable. Positivity of M_{t+1} ensures the absence of arbitrage opportunities. Notice that the vector $\boldsymbol{\gamma} \equiv [\gamma_0\ \gamma_1\ \gamma_2 \cdots \gamma_K]'$ restricts the relationship between the prices of all assets and the priced risk factors to be homogeneous across assets according to the fundamental Euler condition in (2). There is of course considerable latitude in the definition of what are the priced risk factors that enter M_{t+1}.

Either expanding their number ($K \rightarrow \infty$) or by carefully selecting their identity, these will improve the fit of the model.[4] Although it would be interesting to derive (4) from a general equilibrium framework (possibly with production, e.g., as in Fu and Yang, 2012), in this paper we take such a structure of the SDF as a primitive but estimable functional form and discuss instead the identity, the number of the factors, and (in Section 2.2) the number of hidden states affecting the loadings.

At this point, if we assume that the random vector $\boldsymbol{y}_{t+1} \equiv [m_{t+1}\ f_{1,t+1}\ f_{2,t+1}\ \dots\ f_{K,t+1}\ r_{1,t+1}\ r_{2,t+1}\ \dots\ r_{n,t+1}]'$ expanded to include the factors driving the SDF, has a stationary multivariate Gaussian distribution, then Appendix A shows that the asset pricing model

$$E\left[r_{i,t+1}|\Psi_t\right] = -E[m_{t+1}|\Psi_t] - \frac{1}{2}\sigma_m^2 - \frac{1}{2}\sigma_i^2 - \sigma_{i,m} \tag{5}$$

obtains, where $\sigma_m^2 \equiv Var[m_{t+1}], \sigma_i^2 \equiv \frac{1}{2}Var\left[r_{i,t+1}\right]$, and $\sigma_{i,m} \equiv Cov\left[r_{i,t+1}, m_{t+1}\right]$. Because it is well known that the time t price of a riskless, one-period zero coupon bond equals $1/\ E[M_{t+1}|\Psi_t]$, and that as a result, the continuously compounded riskless return is

$$r_t^f = -E[m_{t+1}|\Psi_t] - \frac{1}{2}\sigma_m^2, \tag{6}$$

the model is equivalently re-written as:

$$E\left[r_{i,t+1}|\Psi_t\right] + \frac{1}{2}\sigma_i^2 - r_t^f = -\sigma_{i,m} \tag{7}$$

where $0.5\sigma_i^2$ is a standard Jensen's inequality correction and the driving force behind the risk premium is therefore $-\sigma_{i,m}$, as one would expect: the higher the covariance between asset returns and the SDF (hence, the higher the covariance between asset returns and marginal utility that is high in bad states), the lower the risk premium on the asset.

Appendix A also shows that under a rather standard stationary VAR(P) representation for the process followed by the factors, the full joint model for the SDF and asset returns is:

$$\begin{aligned}
m_{t+1} &= \gamma_0 + \sum_{j=1}^{K} \gamma_j f_{j,t+1} + w_{t+1} \qquad\qquad (w_{t+1} \equiv m_{t+1} - E[m_{t+1}|\Psi_t]) \\
f_{j,t+1} &= \varphi_{j,0} + \sum_{k=1}^{K}\sum_{p=1}^{P} \varphi_{j,k,p} f_{k,t+1-p} + \delta_{j,t+1}, \qquad j = 1, \dots, K \\
r_{i,t+1} &= \underbrace{\left(0.5\sigma_m^2 + 0.5\sigma_i^2 + \sigma_{i,m}\right)}_{\mu_i} + m_{t+1} + v_{i,t+1} = \mu_i + m_{t+1} + v_{i,t+1}, \quad i = 1, \dots, n
\end{aligned} \tag{8}$$

which can also be represented as:

[4] A few special cases may provide meaningful benchmarks. Clearly, when in (4) K = 1 and $f_{1,t+1}$ is the gross return on the wealth process (sometimes called, the market portfolio), then the SDF becomes the standard CAPM. When K = 1 and $f_{1,t+1}$ is the log-consumption growth rate, γ_0 is the subjective rate of discount, and γ_1 is the opposite of the constant coefficient of relative risk aversion), then the SDF reduces to the classical consumption (C)CAPM. De Roon and Szymanowska (2010) test whether commodity futures returns vary cross-sectionally in a systematic way due to differences in consumption risk and they find that at quarterly horizon, the CCAPM explains about 50% of the cross-sectional variation in mean futures returns, while the conditional version explains up to 60%.

$$\begin{bmatrix} 1 & \mathbf{0}_K' & \mathbf{0}_n' \\ \mathbf{0}_K & \boldsymbol{I}_K & \boldsymbol{O}_{K\times n} \\ \mathbf{0}_n & \boldsymbol{O}_{n\times K} & \boldsymbol{I}_n \end{bmatrix} \boldsymbol{y}_{t+1} = \boldsymbol{\mu} + \sum_{l=1}^{P} \begin{bmatrix} 0 & \ddot{\boldsymbol{\gamma}}_l' I_{\{l=0\}} & \mathbf{0}_n' \\ \mathbf{0}_K & \boldsymbol{\Phi}_l & \boldsymbol{O}_{K\times n} \\ \boldsymbol{\iota}_n I_{\{l=0\}} & \boldsymbol{O}_{n\times K} & \boldsymbol{O}_{n\times n} \end{bmatrix} \boldsymbol{y}_{t+1-l} + \boldsymbol{\eta}_{t+1} \text{ or}$$

$$\boldsymbol{A}_0 \boldsymbol{y}_{t+1} = \boldsymbol{\mu} + \sum_{l=1}^{P} \boldsymbol{A}_l\, \boldsymbol{y}_{t+1-l} + \boldsymbol{\eta}_{t+1} \quad with \;\; \boldsymbol{\Phi}_0 = \boldsymbol{O}_{K\times K} \tag{9}$$

where $\boldsymbol{\mu} \equiv \left[\gamma_0\ \varphi_{1,0}\ \varphi_{2,0}\ \dots\ \varphi_{K,0}\ \left(0.5\sigma_m^2(\boldsymbol{\gamma}) + 0.5\sigma_1^2 + \sigma_{1,m}\right) \dots (0.5\sigma_m^2(\boldsymbol{\gamma}) + 0.5\sigma_n^2 + \sigma_{n,m})\right]'$, or

$$\boldsymbol{\mu} \equiv \begin{bmatrix} \gamma_0 \\ \boldsymbol{\varphi}_0 \\ 0.5\sigma_m^2(\boldsymbol{\gamma}) + 0.5 diag_{2:n+1}(\boldsymbol{\Sigma}) + \boldsymbol{\Sigma}\boldsymbol{e}_1 \end{bmatrix} \quad \ddot{\boldsymbol{\gamma}}_l \equiv \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_K \end{bmatrix} \quad \boldsymbol{\Phi}_l \equiv \begin{bmatrix} \varphi_{l,1,1} & \varphi_{l,1,2} & \cdots & \varphi_{l,1,K} \\ \varphi_{l,2,1} & \varphi_{l,2,2} & \cdots & \varphi_{l,2,K} \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{l,K,1} & \varphi_{l,K,2} & \cdots & \varphi_{l,K,K} \end{bmatrix}$$

and $\boldsymbol{\eta}_{t+1} \equiv \left[w_{t+1}\ \delta_{1,t+1}\ \delta_{2,t+1} \dots\ \delta_{K,t+1}\ v_{1,t+1} - w_{t+1} \dots\ v_{n,t+1} - w_{t+1}\right]' \sim IID\ N(0, \boldsymbol{\Sigma})$.

2.2 HMM Factor SDF Models

One of the explanations sometimes adducted for the negative correlation between commodities and other asset classes is the different behavior of stocks, bonds and commodities over the business cycle (see, e.g., Jensen, Mercer and Johnson, 2002; Vrugt, 2003). Indeed, Gorton and Rouwenhorst (2006) find that commodity futures perform well in the early stages of a recession, a time when stock returns generally disappoint while in later stages of recessions, commodity returns fall off, but this is generally a very good time for equities. This suggests that a well-specified model for the SDF ought to account for persistent, good and bad states. Moreover, Bakshi et al. (2014) have noted that conditional pricing models that allow for state dependence in the sensitivity of the stochastic discount factor to the risk factors can often outperform their unconditional counterparts. Following their lead, suppose the SDF has instead a regime switching log-linear structure:

$$M_{t+1}(S_{t+1}) = exp\left(\gamma_{0,S_{t+1}} + \sum_{j=1}^{K} \gamma_{j,S_{t+1}} f_{j,t+1}\right) > 0, \tag{10}$$

where $f_{1,t+1}, f_{2,t+1}, \dots, f_{K,t+1}$ are again K observable, systematic factors driving the way in which all assets are priced and S_{t+1} follows a J-state Markov chain. The state-specific vector $\boldsymbol{\gamma}_{S_{t+1}} \equiv \left[\gamma_{0,S_{t+1}}\ \gamma_{1,S_{t+1}}\ \gamma_{2,S_{t+1}} \dots\ \gamma_{K,S_{t+1}}\right]'$ restricts the within-regime relationship between the prices of all assets and the priced risk factors to be homogeneous across assets according to the standard Euler condition in (1). By construction the same regime shifts affect the mapping between all factors and the SDF and as such these are transmitted to all priced assets or portfolios. Because of the MS structure of (9), the empirical fit of the model to the cross-section of asset returns may be improved not only by expanding or carefully selecting the K factors, but also by defining the number of regimes or the features of the Markov chain driving the shifts in the coefficients loading the factors on the SDF.

Following steps analogous to those in Appendix A and conditioning on both the state S_{t+1}, and the past of observable returns and factors, if the random vector $\boldsymbol{y}_{t+1} \equiv [m_{t+1}\ f_{1,t+1}$

$f_{2,t+1} \ldots f_{K,t+1}\, r_{1,t+1}\, r_{2,t+1} \ldots r_{n,t+1}]'$, with $\Psi_t \equiv (\{\boldsymbol{y}_{t-s}\}_{s\geq 0}, S_t) \subset \Im_t$, has a conditional multivariate Gaussian distribution then the conditional asset pricing model,

$$E\left[r_{i,t+1}|\Psi_t, S_t\right] = \underbrace{\left(0.5\sigma_m^2(S_{t+1}) + 0.5\sigma_i^2 + \sigma_{i,m}(S_{t+1})\right)}_{\mu_i(S_{t+1})} + E[m_{t+1}(S_{t+1})|\Psi_t, S_t]$$
$$= \mu_i(S_{t+1}) + E[m_{t+1}(S_{t+1})|\Psi_t, S_t] \qquad i = 1, \ldots, n \tag{11}$$

obtains. This is a model in which regimes in the SDF are reflected in expected asset returns both directly through the one-step ahead forecast $E[m_{t+1}(S_{t+1})|\Psi_t, S_t]$ and indirectly, via the asset-specific terms $\mu_i(S_{t+1}) \equiv 0.5\sigma_m^2(S_{t+1}) + 0.5\sigma_i^2 + \sigma_{i,m}(S_{t+1})$ that also reflect state-dependent covariances between asset returns and the SDF. Also in this case, even though the forcing, observable state variables, $f_{1,t+1}, f_{2,t+1}, \ldots, f_{K,t+1}$, follow a linear process and may enter linearly the model, the latter becomes non-linear because of the role played by the latent Markov state, S_t, that governs the switches in the parameters appearing in the process of the factors. For instance, the full model may be written as:

$$m_{t+1} = \gamma_{0,S} + \sum_{j=1}^{K} \gamma_{j,S} f_{j,t+1} + w_{t+1} \qquad (w_{t+1} \equiv m_{t+1} - E[m_{t+1}|\Psi_t, S_t])$$
$$f_{j,t+1} = \varphi_{j,0} + \sum_{k=1}^{K}\sum_{p=1}^{P} \varphi_{j,k,p} f_{k,t+1-p} + \delta_{j,t+1}, \qquad j = 1, \ldots, K \tag{12}$$
$$r_{i,t+1} = \mu_i(S_{t+1}) + E[m_{t+1}(S_{t+1})|\Psi_t, S_t] + v_{i,t+1} = \mu_i + m_{t+1} + v_{i,t+1}, \quad i = 1, \ldots, n$$

where the vector of K factors follow a VAR *(P)* process and that can also be represented as:

$$\begin{bmatrix} 1 & \mathbf{0}_K' & \mathbf{0}_n' \\ \mathbf{0}_K & \boldsymbol{I}_K & \boldsymbol{O}_{K\times n} \\ \mathbf{0}_n & \boldsymbol{O}_{n\times K} & \boldsymbol{I}_n \end{bmatrix} \boldsymbol{y}_{t+1} = \boldsymbol{\mu}(S_{t+1}) + \sum_{l=1}^{P} \begin{bmatrix} 0 & \ddot{\boldsymbol{\gamma}}_{l,S_{t+1}}' I_{\{l=0\}} & \mathbf{0}_n' \\ \mathbf{0}_K & \boldsymbol{\Phi}_l & \boldsymbol{O}_{K\times n} \\ \boldsymbol{\iota}_n I_{\{l=0\}} & \boldsymbol{O}_{n\times K} & \boldsymbol{O}_{n\times n} \end{bmatrix} \boldsymbol{y}_{t+1-l} + \boldsymbol{\eta}_{t+1} \tag{13}$$

where $\boldsymbol{\mu} \equiv [\gamma_0\ \varphi_{1,0}\ \varphi_{2,0}\ \ldots\ \varphi_{K,0}\ (0.5\sigma_m^2(\boldsymbol{\gamma}_{S_{t+1}}) + 0.5\sigma_1^2 + \sigma_{1,m}) \ldots (0.5\sigma_m^2(\boldsymbol{\gamma}_{S_{t+1}}) + 0.5\sigma_n^2 + \sigma_{n,m})]'$, or

$$\boldsymbol{\mu}(S_{t+1}) \equiv \begin{bmatrix} \gamma_0 \\ \boldsymbol{\varphi}_0 \\ 0.5\sigma_m^2(\boldsymbol{\gamma}_{S_{t+1}}) + 0.5 diag_{2:n+1}\left(\boldsymbol{\Sigma}(\boldsymbol{\gamma}_{S_{t+1}})\right) + \boldsymbol{\Sigma}(\boldsymbol{\gamma}_{S_{t+1}})\boldsymbol{e}_1 \end{bmatrix}$$
$$\ddot{\boldsymbol{\gamma}}_{l,S_{t+1}} \equiv \begin{bmatrix} \gamma_{1,S_{t+1}} \\ \gamma_{2,S_{t+1}} \\ \vdots \\ \gamma_{K,S_{t+1}} \end{bmatrix} \qquad \boldsymbol{\Phi}_l \equiv \begin{bmatrix} \varphi_{l,1,1} & \varphi_{l,1,2} & \cdots & \varphi_{l,1,K} \\ \varphi_{l,2,1} & \varphi_{l,2,2} & \cdots & \varphi_{l,2,K} \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{l,K,1} & \varphi_{l,K,2} & \cdots & \varphi_{l,K,K} \end{bmatrix}$$

and $\boldsymbol{\eta}_{t+1} \equiv [w_{t+1}\ \delta_{1,t+1}\ \delta_{2,t+1} \ldots \delta_{K,t+1}\ v_{1,t+1} - w_{t+1} \ldots v_{n,t+1} - w_{t+1}]' \sim IID\ N(0, \boldsymbol{\Sigma}_{S_{t+1}})$.[5] For simplicity, we assume a time homogeneous Markov chain, although a literature shows that generalizations may be fruitful (see e.g., Dias et al., 2015). However, in our case, an element of time heterogeneity is already impressed by the fact that the SDF depends on factors characterized by a potentially rich dynamics, while the HMM modelling truly affects

[5] HMM dynamics in the covariance matrix of $\boldsymbol{\eta}_{t+1}$ derives from regimes in the SDF loadings. The assumption of multivariate normal shocks is typical of the literature, see Bae et al. (2014).

not directly the factors, but the loadings with which the factors appear in the SDF, in equation (10).

2.3 Maximum Likelihood Estimation Strategy

Our estimation strategy that is based on the principle of maximum likelihood, under parametric assumptions concerning the joint normal distribution of random vector $\boldsymbol{y}_{t+1} \equiv [m_{t+1}\, f_{1,t+1}\, f_{2,t+1} \ldots f_{K,t+1}\, r_{1,t+1}\, r_{2,t+1} \ldots r_{n,t+1}]'$ and, when appropriate, the Markov state variable S_t.

In the single-state, covariance stationary case, because the inverse of the $(n+K+1)\times(n+K+1)$ matrix $\boldsymbol{A}_0$ obviously exists (its determinant is 1), and the corresponding Jacobian is unity, the log of the joint density function of the sample, conditioned on the initial values of the variables ($\boldsymbol{y}_0$), is given by[6]

$$ln\, f(\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_T; \boldsymbol{\theta}) = -\frac{T}{2}\ln\det(\boldsymbol{\Sigma}) - \frac{1}{2}\sum_{t=1}^{T}\left(\boldsymbol{A}_0\boldsymbol{y}_{t+1} - \boldsymbol{\mu} - \sum_{l=1}^{P}\boldsymbol{A}_l\,\boldsymbol{y}_{t+1-l}\right)'\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{A}_0\boldsymbol{y}_{t+1} - \boldsymbol{\mu} - \sum_{l=1}^{P}\boldsymbol{A}_l\,\boldsymbol{y}_{t+1-l}\right) \qquad (14)$$

(up to an omitted constant term), when the matrices are replaced by the appropriate objects. Maximizing such a log joint density function with respect to the components of $\boldsymbol{\theta} \equiv [\boldsymbol{\gamma}\ \boldsymbol{\varphi}_0\ \{\boldsymbol{\Phi}_l\}_{l=1}^{P}\ \boldsymbol{\Omega}]'$ will deliver the ML estimates of $\boldsymbol{\theta}$. Note that the (conditional) residuals used by the MLE program,

$$\ddot{\boldsymbol{\eta}}_{t+1} \equiv \ddot{\boldsymbol{A}}_0\ddot{\boldsymbol{y}}_{t+1} - \ddot{\boldsymbol{\mu}} - \sum_{l=1}^{P}\ddot{\boldsymbol{A}}_l\,\ddot{\boldsymbol{y}}_{t+1-l} = \begin{bmatrix}\boldsymbol{I}_K & \boldsymbol{O}_{K\times n}\\ \boldsymbol{O}_{n\times K} & \boldsymbol{I}_n\end{bmatrix}\ddot{\boldsymbol{y}}_{t+1} - \begin{bmatrix}\boldsymbol{\varphi}_0\\ 0.5\sigma_m^2(\boldsymbol{\gamma}) + 0.5diag_{2:n+1}(\boldsymbol{\Sigma}) + \boldsymbol{\Sigma}\boldsymbol{e}_1\end{bmatrix} - \sum_{l=1}^{L}\begin{bmatrix}\boldsymbol{\Phi}_l & \boldsymbol{O}_{K\times n}\\ \boldsymbol{O}_{n\times K} & \boldsymbol{O}_{n\times n}\end{bmatrix}\ddot{\boldsymbol{y}}_{t+1-l} \qquad (15)$$

imply the impossibility to concentrate the parameters of the log-likelihood function to separate the estimable elements of $\boldsymbol{\Omega}$ from those appearing in the conditional residuals $\boldsymbol{\eta}_{t+1}$.

In the case of the HMM extension introduced in Section 2.2, because the inverse of the matrix $\boldsymbol{A}_0$ still exists, and the corresponding Jacobian is unity, the log of the joint density function of the sample, conditioned on the initial values of the variables,

$$\ln f(\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_T, \mathrm{S}_0, \mathrm{S}_1, \ldots, \mathrm{S}_T; \boldsymbol{\theta}) = -\frac{T}{2}\ln\det\left(\boldsymbol{\Sigma}(\boldsymbol{\gamma}_{S_{t+1}})\right) - \frac{1}{2}\sum_{t=1}^{T}\left(\boldsymbol{A}_0\boldsymbol{y}_{t+1} - \boldsymbol{\mu}(\boldsymbol{\gamma}_{S_{t+1}}) - \sum_{l=1}^{P}\boldsymbol{A}_l\,\boldsymbol{y}_{t+1-l}\right)'\boldsymbol{\Sigma}(\boldsymbol{\gamma}_{S_{t+1}})^{-1}\left(\boldsymbol{A}_0\boldsymbol{y}_{t+1} - \boldsymbol{\mu}(\boldsymbol{\gamma}_{S_{t+1}}) - \sum_{l=1}^{P}\boldsymbol{A}_l\,\boldsymbol{y}_{t+1-l}\right), \qquad (16)$$

[6] In what follows $\ddot{\boldsymbol{\eta}}_{t+1}$ and $\ddot{\boldsymbol{y}}_{t+1}$ are the same as $\boldsymbol{\eta}_{t+1}$ and $\boldsymbol{y}_{t+1}$ but fail to include the first element of both vectors as the SDF is by construction unobservable.

(up to an omitted constant term) is the same as the log-likelihood function for $\boldsymbol{\eta}_{t+1}$. It turns out that maximizing such a log joint density function with respect to the components of $\boldsymbol{\theta} \equiv \left[\{\boldsymbol{\gamma}_s\}_{s=1}^{J}\ \boldsymbol{\varphi}_0\ \{\boldsymbol{\Phi}_l\}_{l=1}^{P}\ \{\boldsymbol{\Sigma}(\boldsymbol{\gamma}_s)\}_{s=\mathbf{1}}^{J}\right]'$ will deliver the ML estimates of $\boldsymbol{\theta}$. Because the states are latent, an application of the iterative Expectation-Maximization (EM) algorithm in which the maximization is applied to

$$
\begin{aligned}
-\frac{1}{2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{s=1}^{J} \Pr(S_t = s|\Psi_T) \ln\det\big(\boldsymbol{\Sigma}(\boldsymbol{\gamma}_s)\big) + \\
-\frac{1}{2}\sum\nolimits_{t=1}^{T}\sum\nolimits_{s=1}^{J} \Pr(S_t = s|\Psi_T)\Big(\boldsymbol{A}_0\boldsymbol{y}_{t+1} - \boldsymbol{\mu}(\boldsymbol{\gamma}_s) \\
-\sum\nolimits_{l=1}^{P} \boldsymbol{A}_l\,\boldsymbol{y}_{t+1-l}\Big)' \boldsymbol{\Sigma}\big(\boldsymbol{\gamma}_{S_{t+1}}\big)^{-1}\Big(\boldsymbol{A}_0\boldsymbol{y}_{t+1} - \boldsymbol{\mu}(\boldsymbol{\gamma}_s) \\
-\sum\nolimits_{l=1}^{P} \boldsymbol{A}_l\,\boldsymbol{y}_{t+1-l}\Big), \qquad (17)
\end{aligned}
$$

where $\{\Pr(S_t = s|\Psi_T)\}_{t=1}^{T}$ (s = 1, 2, …, J) are the smoothed probabilities derived with the classical, Hamilton-Kim smoothing algorithm (Hamilton, 1994). Maximizing such a smoothed probability-weighted log joint density function delivers ML estimates of $\boldsymbol{\theta}$.

The assumption of normality of returns and finite number of states allows us to employ the Baum–Welch algorithm to estimate $\boldsymbol{\theta}$. The Baum–Welch algorithm is an expectation–maximization (EM) algorithm typically applied to HMM. Given an initial set of parameters $\boldsymbol{\theta}_0$ and the realized series of observations, $\{\boldsymbol{y}_t,\}$, the result of this algorithm always converges to a local maximum of the likelihood function. In the E-step, we compute the expected value of the T+1 latent Markov states (one at each point in the sample) given the observed data and the current, provisional estimates of the parameters. In the M-step, standard ML methods are used to update the unknown parameter estimates using an expanded data matrix with previous expectations as weights.[7] To protect against the perils of multiple, local maxima, as in Bae et al. (2014), we use randomly selected initial parameters and repeat the parameter estimation for each seed, where each new set of initial parameters is drawn from the results of the previous estimation run, assuming a normal distribution with mean equal to the estimate and variance equal to the square of the standard error obtained from the previous run of the algorithm.[8]

[7] Because the EM algorithm needs to store the JT entries of expectations over the latent space for each data pattern, computation time and storage increases exponentially with the number of observations, which makes this algorithm impractical. However, for HMM, the special variant of the EM algorithm referred to as the forward–backward or Baum–Welch algorithm is needed because the model contains a huge number of entries in the joint posterior latent distribution generated by the T+1 latent variables. The Baum-Welch algorithm circumvents the computation of this joint posterior distribution making use of the conditional independencies implied by the model, see Dias et al. (2015) for a discussion.

[8] Random draws of the starting parameter values that violated basic admissibility conditions (e.g., combinations of elements in the $\{\boldsymbol{A}_l\}$ matrices that make the VAR system non-stationary) were rejected.

3 The Data

We use a cross-section of eight portfolios of stocks, Treasuries, corporate bonds and commodity indices as test assets throughout. In particular, we consider the monthly excess returns of these portfolios over the one-month Treasury Bill rate. Given the limited availability of data on commodity-based factors, all series span the sample January 1989 - December 2011.

The stock portfolio is proxied by the CRSP value-weighted index (VW CRSP), which includes all the firms incorporated in the U.S. and listed on the NYSE, AMEX, or NASDAQ, obtained from Fama and French's data library. Treasuries are proxied by the 10-year constant maturity Treasury yields, that is, yields on actively traded non-inflation indexed issues adjusted to constant maturities, taken from the Federal Reserve of St. Louis FRED data repository. The two corporate bond portfolios are proxied by the Moody's seasoned Aaa corporate bond yield portfolio and the Moody's seasoned Baa corporate paper portfolio, respectively, and are obtained from the Federal Reserve of St. Louis FRED data repository. In order to transform yields into returns we use Shiller's (1979) approximation, which defines the holding period yield in terms of yields to maturity. The one-month Treasury Bill return series is obtained from the Ibbotson SBBI Classic Yearbook, as reported in the Fama and French's data library.

Similarly to research by Bae et al. (2014), the four commodity indices included among the test assets are all Standard & Poor's-Goldman Sachs (S&PGS) Spot Commodity Indices, namely the Standard & Poor's-Goldman Sachs Agriculture and Livestock, the S&PGS Precious Metals, the S&PGS Industrial Metals, and the S&PGS Energy indices. The S&P-GSCI spot indices are built using front-end futures to exploit the proximity of traded future prices to spot prices. However, we are aware that an index of commodity spot prices simply tracks the evolution of the spot prices, and ignores all costs associated with the holding of physical commodities (storage, insurance, etc). It is therefore an upper bound on the return that an investor in spot commodities would have earned in real time.

The pricing factors used in this paper can be divided into two groups: macro-based and commodity-specific pricing factors. The macro-based factors are built following the same approach as in Ludvingson and Ng (2009). We start from a rich set of 132 U.S. macroeconomic variables, all measured at monthly frequency. We then remove all financial return series and add three topical series: the Goldman Sachs Financial Index Conditions Index (GS-FCI), the Historical News-Based Policy Index (see Baker, Bloom, Davis, 2012) and the Liquidity Factor of Pastor and Stambaugh (2003).[9] We summarize the covariance

[9] The Goldman Sachs Financial Stability Index is a weighted sum of a short-term bond yield, a long-term corporate yield, the exchange rate and a stock variable. The Historical News-Based Policy Index is an index of economic policy uncertainty for the U.S., constructed from three types of underlying components: newspaper coverage of policy-related economic uncertainty, the number of federal tax code provisions set to expire in future years, and disagreement among economic forecasters as a proxy for uncertainty.

structure of this set of variables using principal component analysis as in Ludvingson and Ng (2009). We extract three/five orthogonal factors that summarize 35% and 44% of the variance, respectively. To gain an insight on the economic meaning of these factors we investigate the sign of the factor loadings and the R-squares from univariate regressions of each of the factors on each of the baseline macroeconomic variables (see Table 1).

The first principal component (F1) is a business cycle, pro-cyclical factor, the value of which increases with industrial production, help wanted, and employment growth and decreases with changes in the unemployment rate. The second principal component (F2) can be defined as an inflation factor that positively loads with relatively high R-squares on most types of price indices (including consumers', producers', and personal consumption expenditure deflators). The third principal component (F3) can be defined as an inventory and new orders factor.

The three commodity factors are the hedging pressure (HP) factor, the basis factor and the momentum factor (Daskalaki et al., 2014).[10] The HP factor is defined as the difference between the return of a portfolio of commodity futures with positive hedging pressure and the return of a portfolio with negative hedging pressure, where the hedging pressure for commodity *i* at time *t* is defined as the number of short hedging positions minus the number of long hedging positions divided by the total number of hedgers in the considered commodity market. The basis factor is the difference between the return of a portfolio of commodity futures with positive basis and the return of a portfolio of commodity futures with negative basis.[11] The momentum factor is given by the difference between the return of a portfolio of commodity futures with positive prior 12-month return and the return of a portfolio of futures with negative prior 12-month return. The HP factor is motivated by the hedging pressure hypothesis of Keynes and Hicks: according to this hypothesis, the futures market provides a risk transfer mechanism whereby risk-adverse speculators demand a compensation for taking the price risk of hedgers. The basis factor is motivated by the theory of storage which predicts that the sign of the basis depends on the convenience yield: high convenience yield is associated with a positive basis and vice versa. The momentum factor is documented in Gorton et al. (2013) and is linked to the time series variation of inventories. In a theoretical perspective, positive hedging pressure, positive

Pastor-Stambaugh's Liquidity Factor is a cross-sectional average of individual stock liquidity measures. Each stock liquidity in a given month represents the effect that a given volume on day t has on the return for the day t+1.This measure of market-wide liquidity, hence, captures a dimension of liquidity associated with the strength of volume-related return reversals and is based on the idea that order flows induce greater return reversals when liquidity is lower.

[10] For the construction of these factors, 22 commodity futures contracts during January 1989-December 2013 obtained from Bloomberg and the number of long/short hedgers provided by the Commodity Futures Trading Commission are used. We thank George Skiadopoulos for sharing the series with us.

[11] The basis for commodity *i* at time *t* is given by$Basis_{i,t} \equiv (F_{i,t} - F_{i,t+1})/F_{i,t}$, where $F_{i,t}$ indicates the price of the nearest futures contract on commodity *i* and $F_{i,t+1}$ represents the price of the next nearest futures contract.

basis and high prior futures returns (high momentum) should be associated with positive futures returns and vice versa.

Table 2 shows summary statistics. The normality hypothesis is rejected according to the Jarque-Bera (JB) test for all series, including macro factors. The three macro factors are characterized by a mean equal to zero and by a considerable volatility. This finding is explained by the fact that all factors are characterized by large and positive excess kurtosis and non-zero (and highly statistically significant, based on unreported tests) skewness. The evidence of non-normality extends to the commodity factors and all the portfolios considered, even though bonds are characterized by negative excess kurtosis and essentially symmetric empirical distributions, and the rejection of the null of normality tends to be relatively weak, with p-values between 1 and 5 percent. Of course, the HMM methods developed in Section 2 can been seen as ways to capture and forecast such pervasive departures from normality, as discussed in Buckley et al. (2008) and Dias et al. (2015). Interestingly, commodities show high return volatility not always associated to a high mean return: in fact, the Agriculture & Livestock class displays a negative mean return despite a standard deviation close to that of stocks.

Table 2 also shows the correlations among the series. Interestingly, F2 is the pricing factor showing the highest correlations with the different test assets and this seems to imply that inflation tends to simultaneously correlate with commodity returns, which is consistent with some of the literature on commodities as inflation hedge (see, e.g., Erb and Harvey, 2006). The correlations among the test assets and the remaining pricing factors are generally low with the exception of Industry Metals which positively and significantly correlate with Precious Metals and Agriculture and Livestock. However, the range of correlations between pairs of commodities indices is wide, spanning from 0.16 between agricultural and energy commodities to 0.31 between agricultural and industrial commodities.

4 Empirical Results for Macro-Based Linear SDF Models

4.1 Model Selection and Empirical Estimates

In Table 3, we compare the values of three different information criteria—the Akaike (AIC), Schwarz (SIC), and Hannan-Quinn (HQC) criteria—to select the model specification that most likely represents the underlying but unknown data generating process. As discussed by Dias et al. (2015), the standard model selection approach when using MLE is by performing likelihood ratio tests across nested models, here the relevant tests are between models with J – 1 and J regimes. However, in the context of HMM, this approach is problematic because the null hypothesis under test is defined on the boundary of the parameter space, and consequently the regularity condition on the asymptotic properties of the MLE are not valid. The basic principle behind information criteria is parsimony, which results from the trade-off between model fit and model complexity.

The table is organized around three panels. In the first part of the table, we entertain single-state models, with constant SDF, three factors (either macro-based or commodity-specific), five factors (macro principal components), and eight (when all factors available are simultaneously specified). In all these cases, the SDF has a simple log-linear structure. What further differentiates the competing models is the number of lags in the factor HMM VAR, also called MSVAR because we use a Markov switching VAR model and the structure of the factor MSVAR (full MSVAR, block MSVAR, diagonal MSVAR). The concept of structure of the factor MSVAR we are referring to the $\mathbf{\Phi}_l$ matrices in equations (9) and (13), l = 1, 2, … P. In case of full MSVAR, $\mathbf{\Phi}_l$ is a full matrix (each factor depends on its own lags, and on past values of both the other pricing factors and of the test assets); in case of block MSVAR, $\mathbf{\Phi}_l$ is a block matrix (each factor depends on its own lags and on past values of the other pricing factors); in case of a diagonal MSVAR, $\mathbf{\Phi}_l$ it is a diagonal matrix (each factor depends only on its own past values). The second and third panels show the same models estimated to display two and three regimes, respectively.

The rationale behind information criteria is to provide a measure of information that strikes a balance between goodness of fit and parsimonious specification of the model. Consequently, they select the most parsimonious correct model.[12] The two less parsimonious criteria (i.e., AIC and HQC) select the two-state, one lag, full factor VAR based on commodity factors only. This is a relatively rich model with a satisfactory saturation ratio (almost 37).[13] The second model selected by the SIC is a single-state, o diagonal factor VAR(1) based on commodity factors only, which is very a parsimonious model with only 10 parameters and a saturation ratio equal to 302.5.

Therefore, a formal model specification search similar to Guidolin and Timmermann's (2006) leads to a preference for a mix of relatively "light" and single-state models with diagonal factor structure and of richer, Hidden Markov, full VAR models in which good and bad states are reflected in the dynamics of a nonlinear SDF. Interestingly, the different information criteria leave some uncertainty only concerning the correct specification of the number of regimes (SIC favoring single-state models vs. AIC and HQC favoring two-state models), but leave no doubt as to the fact that not only the three commodity-specific factors (hedging pressure, basis, and momentum) can price the cross-section of commodity spot returns, but also stocks and bonds. Understandably however, AIC and HQC do penalize

[12] All the information criteria are based on -2 times the average log likelihood function adjusted by a penalty function, $\varphi(T)$, multiplied by the number of estimated parameters. In the Akaike information criterion, $\varphi(T) = 2$; in the Schwarz criterion , $\varphi(T) = \ln T$; in the Hannan-Quinn information criterion, $\varphi(T) = 2\ln(lnT)$.

[13] The saturation ratio is defined as the ratio between the number of observations and the number of estimated parameters. A saturation ratio in excess of 20 tends to be recommended in applied econometrics.

models enough that they select HMM SDFs with three factors only but in which the factors follow a rich, full VAR dynamics. Interestingly, when three factors are selected, these are never the macro ones.

To give an idea of the resulting estimates, Tables 4 and 5 show the empirical results for the two winner models. Table 4 concerns the lean, single-state, commodity factor-only diagonal VAR(1) model. In particular, the tables shows the AR(1) coefficients and the intercept for each of the factors, the loadings of the SDF on each of the factors (the γs) and the intercept γ_0, and the residual correlation matrix of the factors and test assets. Interestingly, this simple model is governed by a weak dynamics, in the sense that the persistence of each of the factors is mild and precisely estimated only in the case of momentum (the p-value is 0.099 but the coefficient is 0.087 only). On the contrary, the SDF strongly depends on the three factors and the corresponding loadings are all accurately estimated with zero p-values. Interestingly, the estimated SDF loads positively on the HP and basis factors and negatively on momentum. This makes some intuitive sense: the SDF tends to be relatively low during economic booms and bull markets, when momentum in many asset classes (and especially commodities) tends to be strong and hedging pressures relatively modest because overshadowed by speculative activity affecting all "financialized" markets.

Table 5 shows instead parameter estimates for the more complex, two-state full VAR(1) model with commodity factors only. The last column of the table shows that a number of coefficients are strongly switching across Markov states, in particular the intercepts of the VAR model governing the dynamics of the factors (these are systematically higher in regime 2), the persistence of the HP factor (that changes sign, from 0.20 to -0.05 going from regime 1 to regime 2), and especially the loadings of the SDF on the HP and momentum factors (the former higher and the latter lower, in regime 1 vs. regime 2). However, the loading of the SDF on the HP factor are never accurately estimated, in any of the regimes. Also the (co)variances of the system residuals are strongly regime switching (as indicated by the appearance of two sets of estimates in Table 5, separated by a bar symbol in the case of the variances on the main diagonal): residuals in regime 1 tends to be more volatile than regime 2 is and covariances are also larger (see Buckley et al., 2008, for a discussion). In fact, as typical of much empirical finance literature (see they survey in Guidolin, 2011), regime 1 can be characterized as a bear, recession state in which the SDF tends to be large and bond yields and expected risky returns low, while regime 2 is a bull, expansionary state with opposite features. Consistently with the bulk of the empirical asset pricing literature (see Dias et al., 2015, and references therein), the good, bull regime is highly persistent with a "stayer" probability of 0.91 and an average duration of 11 months, while the recessionary, bear state is less persistent with a probability of staying in this regime of 0.81 and an average duration of 5 months. Overall, these estimates of the Markov transition matrix imply long-run, ergodic probabilities of 0.69 and 0.31, respectively, for regimes 2 and 1.[14]

An unreported table has also tabulated results for two macro-based SDF models, corresponding to the commodity-based models in Tables 4 and 5. We know from Table 3 that these models are hardly selected by the information criteria. An analysis of the tabulated results easily reveals why: even though the macro factors are characterized by remarkable persistence and rich dynamics (this is particularly the case of factors 1 and 3), and the loadings of the SDF are (in both regimes, when these are specified) precisely estimated, such loadings tend to be extremely small and this turns out—given that the volatility of the first three principal components is comparable to those of portfolio returns—to be insufficient to transmit sufficient intertemporal variation to the SDF. Of course, this represent a common problem with macro-based SDFs in all of the asset pricing literature (see e.g., the review in Cochrane, 2005). Surprisingly though, this lack of sufficient variation tends to also concern HMM macro-based SDFs, contrary to what has been found elsewhere, although on different data set and using alternative econometric techniques (see the review in Guidolin, 2012).

4.2 Observed and Implied Moments

In this subsection we assess to what extent a range of competing SDFs are able to match the empirical mean, standard deviation, skewness, excess kurtosis, and pair-wise cross asset correlations, with special emphasis on results concerning commodities. Specifically, we check whether each model-implied moment falls in the 90% confidence band built around the sample estimate. If this happens, we consider this moment to be matched by the model. The moments implied a given SDF are computed by simulations, using 20,000 trials and as such they can be taken to represent population moments implied by each given model. In particular, we compare four of the competing model specifications discussed in Section 4.2 and covered by Table 3, namely, a single-state diagonal VAR(1) including only commodity factors, a single-state diagonal VAR(1) including only macro factors, a single-state diagonal VAR(1) including both commodity and macro factors, and a full two-state MSVAR(1) including only commodity factors.

Table 6 shows that while all competing models are able to match almost all the empirical volatility estimates, with only minor exceptions concerning the standard deviations of the factors (both macro and commodity related ones), the outlook is much more heterogeneous as far as the mean and the skewness coefficients are concerned. In particular, the single-state VAR that includes only the macro factors performs really poorly at replicating both the sample mean and skewness. Indeed, this SDF is just able to match the

[14] Unreported plots suggest that regime 1 may correspond to either the 2007-2011 financial crisis or to the financialization of commodities (or both). This is in line with the interpretation of state 1 as the crisis state with a different SDF. Indeed we find that the implied SDF is on average higher in crisis time.

mean of two asset classes out of eight (equity and energy commodities), and none of the sample skewness coefficients. All of the other model specifications are able to match the mean of the majority of the asset classes, with the exclusion of bonds. Instead, the two-state HMM SDF based on commodity factors is weaker than the other two at reproducing the right level of skewness. Finally, none of the models seem to be able to replicate the empirical excess kurtosis (with some exceptions for the regime switching SDF, as one would expect on the basis of the literature, see e.g., Dias et al, 2015).

As far as pair-wise correlations are concerned, an unreported table (available on request from the authors) shows that the single-state diagonal VAR(1) model including only commodity factors is the best performer, as it is able to replicate 43 correlations out of 55, with a success ratio of 78%. However, also the single-state diagonal VAR including all factors and the two-state HMM that includes only commodity factors, perform well, as they match 76% and 71% of the correlations, respectively. Conversely, the single-state diagonal VAR model based only on macro factors is able to correctly replicate only 36% of the empirical correlations, confirming its poor fit, as already observed in the case of sample mean and skewness.

Interestingly, the two-state SDF is not the best model at reproducing correlations. This is quite surprising considering that, from a theoretical point of view, given the time-varying correlation between asset classes, a time-varying model may be often expected to deliver better results. This should be particularly true for commodities, given the change in their relationships with traditional assets classes that has (allegedly) occurred over the last decade (the so called "financialization" of commodities). For example, Lombardi and Ravazzolo (2013) find that a bivariate Bayesian dynamic conditional correlation model, which can account for time variation in the correlation patterns, produces statistically more accurate density forecasts for equity and commodity prices and gives large economic gains in an asset allocation exercise, relative to a benchmark random walk model. However, our two-state HMM SDF is the only model able to match the excess kurtosis of at least some of the assets, which is instead to be expected.

Figures 1 through 4 focus on the pair-wise correlations that are most useful for the scope of our analysis. In particular, we are interested in within-asset class correlations (in particular intra-commodity correlations) and across-asset class correlations (in particular correlations between commodities and traditional asset classes). Figure 1 shows that the single-state VAR model including only commodity factors is able to match 25 out of 28 cross asset correlations, with a rate of success of 89%. In particular, it only fails to match the sample correlations of the 10-year Treasury – Aaa Corporate Bonds, 10-year Treasury – Baa Corporate Bonds, and Aaa Corporate Bonds – Baa Corporate Bonds pairs. Conversely, Figure 2 shows that a single-state diagonal VAR model including only macro factors performs poorly in terms of matching the correlations, with a modest success rate of 32%. In particular, it fails to match the correlation between commodities and traditional asset classes, being able to replicate only 2 of them out of 16 (namely, Industrial Metals – CRSP Equity and Energy – CRSP Equity pair-wise correlations). On the contrary, the correlations implied by this model match sample observed intra-bond correlations, which are instead not correctly reproduced by the single-state VAR model including only commodity factors.

Figure 3 clearly shows that the single-state VAR including all factors is the model which better replicates sample observed correlations. Indeed, model implied correlations do not match the empirical ones only in two cases, namely, Agriculture and Livestock – Baa Corporate Bonds and 10-year Treasury – Baa Corporate Bonds pair-wise correlations. Finally, Figure 4 shows that a two-regime model including only commodity factors is able to match 22 out of 28 pair-wise correlations, with a success rate of 79%. Compared to the single-state VAR based on only commodity factors, the two-state model is able to match 10-year Treasury – Aaa Corporate Bonds and 10-year Treasury – Baa correlations. On the contrary, it fails to match Aaa Corporate Bonds – CRSP Equities and Baa Corporate Bonds – CRSP Equities correlations. In addition, it is not able to replicate the pair-wise correlations between Precious Metals and all the fixed income portfolios under examination.

In conclusion, among the commodity factor models, the single-state diagonal VAR model, selected by the SIC criterion (as discussed in Section 4.2) seems to be better than the more richly parametrized two-regime full VAR HMM models at matching empirical correlations. Nonetheless, the two-regime model delivers better results as far as the replication of intra-bond correlations are concerned. In addition, the SDF that projects on all factors, despite not being selected by any information criteria, shows a high ability to replicate sample correlations. In particular, even though this model does not have the highest rate of success at replicating the total number of correlations (76% vs. 78% of the single-state commodity factor model), it is the one matching the highest number of cross asset correlations, which are the most revealing for the scope of our analysis and that directly speak as to the issue of whether commodities can actually be used in portfolio management because of their de-correlation properties.

4.3 Commodity-Specific Factors and the Issue of Segmentation

In this section, we discuss the results presented above in light of the debate concerning whether or not commodities constitute an asset class segmented from more traditional investments such as stocks and bonds. This question is particularly relevant to portfolio allocation, as the answer determinates whether or not commodities represent a good source of diversification for investors. A recent literature (see, e.g., Büyüksahin and Robe, 2013; Miffre, 2011; Tang and Xiong, 2012) has challenged the alleged diversification ability of commodities, arguing that in the last two decades the financial sector has gained influence relative to the real sector in determining the price dynamics in the commodity markets (the already mentioned financialization of commodities). A subset of this literature focuses on the return correlations between commodity futures and other financial assets (e.g., stocks or bonds) and the co-movement among commodities futures prices themselves to determine the extent to which commodity markets have been influenced by financialization. This part of the literature is generally motivated by the view of financial investors as index investors who focus on the strategic portfolio allocation between commodities and other asset classes and thus tend to trade in and out of many commodities at the same time. This argument predicts that financialization causes price co-movements between commodities futures and other major asset classes to increase, more for indexed commodities than for those not in the indices.[15]

Our results show clear evidences of segmentation of commodities from the rest of the asset class universe, as commodity-specific pricing factors are necessary for an SDF model to be able to correctly replicate both intra-commodity correlations and especially the pair-wise correlations between commodities and stocks and bonds. On the other hand, a model including only commodity factors struggles to reproduce intra-bond correlations. When both macro and commodity-specific pricing factors are taken into account, the model-implied correlations match sample observed intra asset class and cross asset correlations in 93% of the cases. Consequently, it seems that commodity-specific factors and macro factors are both needed if one aims at correctly describing the joint behavior of traditional asset classes and commodities.

Our results are in line with some evidence presented in earlier literature. For example, Gorton and Rouwenhorst (2006) report that while the risk premium on commodity futures is essentially the same as equities, commodity futures returns are negatively correlated with equity and bond returns and that this negative correlations tend to increase with the holding period. These findings suggest that commodity futures are effective means for diversifying equity and bond portfolios and that the diversification benefits of commodity futures tend to be larger at longer horizons. However, it is interesting to notice that our results show evidence of a greater segmentation between commodities and bonds rather than between commodities and equity. Indeed, as showed by Figure 1 and 4, intra-bond correlations are difficult to be replicated using only commodity factors, while almost all equity-related empirical correlations are matched by all SDF models with the only exception of the SDF that simply includes macro factors.

5. Conclusions

In this paper we have investigated whether it is possible to find a SDF that jointly prices the cross-section of eight portfolios of stocks, Treasuries, corporate bonds and commodities and

[15] The financialization process exerts different effects: first, standard models of demand and supply are no longer sufficient to explain commodity returns (see e.g., Le Pen and Sévi, 2013); second, the diversification benefits of commodities may be put into question.

replicates the empirical correlations among these assets. Our notion of "pricing" consists of being able to match the first four moments of the empirical distribution of the returns on stocks, bonds, and commodities, with particular emphasis on correlations between commodities considered in pairs as well between commodities and other asset classes. Our SDFs are further extended to include latent regime shifts governed by an ergodic and irreducible Markov chain.

We find that when commodities are considered, commodity-based factors should be preferred over the macro-based ones for pricing purposes. Observed correlations for commodities are perfectly matched by a parsimonious, single state, diagonal factor VAR model where only three commodity-based factors enter the SDF. Finally, there is only mild evidence of a need of additional parametric flexibility in the SDF in the form of hidden Markov regimes for pricing purposes, while regimes are occasionally beneficial as far as matching empirical correlations is concerned.

Our findings are far from being a definitive answer to the asset pricing questions introduced in this paper. First, we have tested only two families of SDFs: macro-based and commodity factor-based models*.* The range of SDF*s* to be tested is much wider. For instance, it would be interesting to investigate whether the factor mimicking portfolios which have been found to explain the cross-section of stock returns are also able to characterize the cross-section of commodity returns. Second, the macro-based factors used in this thesis are the first three PCs extracted from a rich set of 112 US macro-variables. Because commodities are global assets (in the sense that, differently from equities and bonds, it is not possible to distinguish, e.g., European and American commodities) it could be argued that global macro-variables should be considered as factors. It would be interesting to repeat our analysis staring from a set of global macro-variables and check whether our results also hold in this richer framework.

Finally, other commodity-specific factors have been used in the literature. Two examples are given by the commodity liquidity risk factor (see, e.g., Daskalaki et al., 2014) and the aggregate open interest factor (see Hong and Yogo, 2012). The first factor measures the illiquidity in commodity futures markets; the second one measures the open interest in commodity futures markets.

References

Alizadeh, A.H., Nomikos, N. K. , Pouliasis, P. K. (2008). "A Markov Regime Switching Approach for Hedging Energy Commodities". *Journal of Banking and Finance* 32, 1970-1983.

Asness, C. A., Moskowitz, T. J., Pedersen, L.H. (2013). "Value and Momentum 'Everywhere'". *Journal of Finance* 68, 929-985*.*

Ang, A., Bekaert, G. (2002). "Regime Switches in Interest Rates". *Journal of Business and Economic Statistics* 20, 163-82.

Bae, G., Kim, W., Mulvey, J. M. (2014). “Dynamic Asset Allocation for Varied Financial Markets under Regime Switching Framework.” *European Journal of Operational Research* 234, 450-458.

Baker, S.R., Bloom, N. Davis, S.J. (2012). “Has Economic Policy Uncertainty Hampered the Recovery?” Chicago Booth Research Paper No. 12-06.

Bakshi, G., Gao, X., Rossi, A. (2014). “A Better Specified Model to Explain the Cross-Section and Time series of Commodity Returns”. Working Paper, University of Maryland.

Basak, S. Pavlova, A. (2013). “Asset Prices and Institutional Investors”. *American Economic Review* 103, 1728-58.

Basu, D., Miffre, J. (2013). “Capturing the Risk Premium of Commodity Futures: the Role of Hedging Pressure”. *Journal of Banking and Finance* 37, 2652-2664.

Buckley, I., Saunders, D., Seco, L. (2008). “Portfolio Optimization when Asset Returns Have the Gaussian Mixture Distribution.” *European Journal of Operational Research* 185, 1434-1461.

Büyükşahin, B. Robe, M. A. (2013). “Speculators, Commodities and Cross-Market Linkages”. *Journal of International Money and Finance* 42, 37-80.

Cochrane, J. H. (2005). *Asset Pricing.* Revised Edition, Princeton University Press, Princeton, N.J.

Daskalaki,C., Skiadopoulos, G. (2011). “Should Investors Include Commodities in their Portfolios after All? New Evidence”. *Journal of Banking and Finance* 36, 2260-2273.

Daskalaki, C; Kostakis, A., Skiadopoulos, G. (2014). “Are there Common Factors in Commodity Futures Returns?” *Journal of Banking and Finance* 40, 346-363.

De Roon, F. A., Nijman T. E., Veld, C (2000) “Hedging Pressure Effects in Futures Markets”. *Journal of Finance* 55, 1437-1456.

De Roon, F.A., Szymanowska, M., (2010). “The Cross Section of Commodity Futures Returns”. Working Paper. Erasmus University.

Dias, J., Vermunt, J., Ramos, S. (2015). “Clustering Financial Time Series: New Insights from an Extended Hidden Markov Model.” *European Journal of Operational Research* 243, 852-864.

Dhume, D. (2011). “Using Durable Consumption Risk to Explain Commodities Returns”. Working Paper, Federal Reserve Board.

Fu, J., Yang, H. (2012). “Equilibrium Approach of Asset Pricing under Lévy Process.” *European Journal of Operational Research* 223, 701-708.

Gorton, G. Rouwenhorst, K.G. (2006) “Facts and Fantasies about Commodity Futures”. *Financial Analyst Journal* 62, 47-68*.*

Gorton, G.B, Hayashi, F., Rouwenhorst, K.G. (2013). “The Fundamentals of Commodity Futures Returns”. *Review of Finance*, European Finance Association, Vol. 17 (1), pp. 35-105.

Guidolin, M. Timmermann, A. (2006). “An Econometric Model of Nonlinear Dynamics in the Joint Distribution of Stocks and Bonds Returns”. *Journal of Applied Econometrics* 21, 1-22.

Guidolin, M. (2011). “Markov Switching Models in Empirical Finance”. *Advances in econometrics* 27, 1-86

Guidolin. M. (2012). “Markov Switching in Portfolio Choice and Asset Pricing Models: A survey”. *Advances in Econometrics* 27, 87-178

Hamilton, J. D. (1994). *Time series analysis*. Princeton University Press, Princeton, NJ.

Hong, H., Yogo, M. (2012) "What does Futures Markets Interest Tell Us About the Macroeconomy and Asset Prices"? *Journal of Financial Economics* 105, 473-490.

Jensen, G. R., Mercer, J. M. Johnson, R. R. (2002). "Tactical Asset Allocation and Commodity Futures". *Journal of Portfolio Management*, pp. 100-111.

Kat, H. M., Oomen, R.C (2007). "What Every Investor Should Know About Commodities. Part II: Multivariate Return Analysis". *Journal of Investment Management* 5, pp. 40-64.

Koijen, R. S. J., Moskowitz, T. J., Pedersen, L. H., Vrugt, E. B. (2013). "Carry". Fama-Miller Center Working Paper.

Lee, H., Yoder, J. (2007). "A Bivariate Markov Regime Switching GARCH Approach to Estimate Time Varying Minimum Variance Hedge Ratios". *Applied Economics* 39, 1253-1265.

Lombardi, M., Ravazzolo, F. (2013). "On the Correlation Between Commodity and Equity Returns: Implications for Portfolio Allocations". BIS Working Paper 420.

Ludvigson, S.C., Ng, S. (2009). "Macro factors in Bond Risk Premia". *Review of Financial Studies* 22, 5027-5067.

Le Pen,Y. and Sévi, B. (2013). "Futures Trading and the Excess Comovements of Commodity Prices". AMSE Working Papers, 1301. Aix en Marseille School of Economics, Marseille.

Marroquìn Martinez , N., Moreno, M. (2013). "Optimizing Bounds on Security Prices in Incomplete Markets. Does Stochastic Volatility Specification Matter?" *European Journal of Operational Research*, 225, 429-442.

Miffre, J., Georgios Rallis (2007). "Momentum Strategies in Commodity Futures Markets". *Journal of Banking and Finance* 31, 1863-1886.

Miffre, J. (2011). "Long-short Commodity Investing: Implication for Portfolio Risk and Market Regulation". Working Paper, EDHEC- Risk Institute.

Pastor, L., Stambaugh, R. L. (2003). "Liquidity Risk and Expected Stock Returns". *Journal of Political Economy* 111, 642-685.

Shiller, R., J. (1979). "The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure". *Journal of Political Economy* 87, 1190-1219.

Singleton, K. (2012). "Investor Flows and the 2008 Boom/Bust in Oil Prices". Working paper, Stanford University.

Szymanowska, M., de Roon, F.A., Nijman, T.E., Van den Goorbergh, R. (2012). "An Anatomy of Commodity Futures Risk Premia". *Journal of Finance* 69, 453-482.

Tang, K. and Xiong, W. (2012). "Index Investments and the Financialization of Commodities". *Financial Analyst Journal* 68, 54-74.

Vrugt, E.B. (2003). "Tactical Commodity Strategies in Reality: An Economic Theory Approach". ABP Working Paper.

Yang, F. (2013). "Investment Shocks and the Commodity Basis Spread". *Journal of Financial Economics*, 110, 164-184.

Table 1

Factor Loadings and their Interpretation

The following table refers to the principal component decomposition of the 112 U.S. macro-variables presented in section 2. It shows the R^2 statistics from univariate regression of each of the macroeconomic variables on each factor and the corresponding factor loadings.

Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
	Univariate Regression R-square					Factor Loadings				
PI	0.0982	0.0232	0.0032	0.0667	0.0054	0.0716	-0.0495	0.0221	0.1138	0.0366
PI_LESS_TRANSFERS	0.3039	0.0394	0.0022	0.0349	0.0057	0.1260	-0.0646	-0.0183	0.0823	0.0375
REAL_CONSUMPTION	0.0773	0.0074	0.1064	0.0258	0.2097	0.0635	0.0279	0.1273	0.0707	-0.2280
MANUFACTURIN__TRADE_SAL	0.2137	0.0002	0.2275	0.0009	0.0287	0.1056	0.0045	0.1861	0.0133	-0.0843
RETAIL_SALES	0.0449	0.0183	0.1783	0.0724	0.2153	0.0484	0.0440	0.1647	0.1186	-0.2310
IP__TOTAL	0.6524	0.0056	0.0934	0.1265	0.0020	0.1845	-0.0243	0.1193	-0.1567	0.0223
IP__PRODUCTS	0.6010	0.0004	0.0556	0.2377	0.0035	0.1771	-0.0066	0.0920	-0.2149	0.0293
IP__FINAL_PROD	0.4817	0.0033	0.0570	0.3266	0.0016	0.1586	-0.0187	0.0931	-0.2518	0.0200
IP__DURABLE_CONSUMER_GOO	0.2483	0.0037	0.1067	0.1897	0.0552	0.1138	-0.0198	0.1274	-0.1919	0.1170
IP___CONSUMER_GOODS	0.2564	0.0012	0.0580	0.4442	0.0025	0.1157	-0.0111	0.0940	-0.2937	0.0251
IP__CONS_NONDBLE	0.0645	0.0000	0.0009	0.2778	0.0196	0.0580	-0.0009	0.0115	-0.2323	-0.0697
IP__BUS_EQPT	0.4823	0.0036	0.0273	0.0673	0.0000	0.1587	-0.0196	0.0645	-0.1144	-0.0023
IP__MATERIALS	0.4820	0.0128	0.1013	0.0233	0.0006	0.1586	-0.0367	0.1242	-0.0672	0.0117
IP__DURABLE_GOODS_MATERI	0.6768	0.0085	0.0403	0.0112	0.0234	0.1879	0.0300	0.0783	-0.0466	0.0762
IP__NONDURABLE_GOODS_MAT	0.1755	0.0045	0.1479	0.0008	0.0002	0.0957	-0.0218	0.1500	0.0122	0.0076
IP__MANUFACTURING	0.7247	0.0000	0.1172	0.0512	0.0157	0.1945	-0.0011	0.1335	-0.0997	0.0623
IP__RESIDENTIAL_UTILITIE	0.0017	0.0000	0.0309	0.2560	0.0305	-0.0095	0.0021	-0.0685	-0.2230	-0.0870
IP__FUELS	0.0076	0.0139	0.0531	0.0067	0.0016	0.0199	-0.0384	0.0899	0.0359	-0.0201
NAPM_PRODUCTION_INDEX	0.5512	0.0056	0.0243	0.0000	0.0109	0.1696	0.0243	-0.0608	0.0001	0.0519
CAPACITY_UTILIZATION	0.5706	0.0000	0.1634	0.0799	0.0457	0.1726	0.0007	0.1577	-0.1246	0.1065
INDEX_OF_HELP_WANTED_ADV	0.1753	0.0076	0.0018	0.0041	0.1185	0.0957	0.0283	0.0166	-0.0281	0.1714
RATIO_OF_HELP_WANTED_ADS	0.2726	0.0040	0.0002	0.0157	0.1047	0.1193	0.0207	-0.0050	-0.0552	0.1611
CIVILIAN_LABOR_FORCE__EM	0.2625	0.0056	0.0262	0.0698	0.0189	0.1171	0.0244	-0.0632	0.1164	0.0685
CIVILIAN_LABOR_FORCE_01	0.2381	0.0018	0.0223	0.0726	0.0178	0.1115	0.0139	-0.0583	0.1187	0.0663
UNEMPLOYMENT_RATE__ALL_W	0.3838	0.0000	0.0294	0.0051	0.0027	-0.1415	-0.0015	0.0669	0.0316	-0.0259
UNEMP_BY_DURATION__AVERA	0.0333	0.0000	0.0406	0.0255	0.1892	-0.0417	-0.0005	0.0786	0.0703	0.2166
UNEMPLOY_BY_DURATION__PE	0.0037	0.0041	0.0005	0.0009	0.1000	-0.0139	-0.0208	0.0083	-0.0130	-0.1574
UNEMPLOY_BY_DURATION_01	0.0867	0.0004	0.0047	0.0013	0.0033	-0.0673	0.0065	-0.0268	0.0159	-0.0286
UNEMPLOY_BY_DURATION_02	0.2988	0.0014	0.0421	0.0237	0.0564	-0.1249	-0.0121	0.0800	0.0678	0.1182
UNEMPLOY_BY_DURATION_03	0.1447	0.0042	0.0011	0.0298	0.0207	-0.0869	-0.0211	0.0131	0.0761	0.0717
UNEMPLOY_BY_DURATION_04	0.1823	0.0002	0.0595	0.0023	0.0346	-0.0975	0.0043	0.0952	0.0211	0.0926
INITIAL_CLAIMS_FOR_UNEMP	0.1350	0.0009	0.1224	0.0010	0.0130	-0.0839	0.0097	-0.1365	-0.0140	-0.0568
EMPLOYEES_ON_NONFARM_PAY	0.7192	0.0042	0.0521	0.0535	0.0215	0.1937	0.0210	-0.0891	0.1019	-0.0730
EMPLOYEES_ON_NONFARM_01	0.7492	0.0003	0.0624	0.0417	0.0070	0.1977	-0.0058	-0.0975	0.0900	-0.0416
EMPLOYEES_ON_NONFARM_03	0.0086	0.0404	0.0628	0.2308	0.0038	0.0212	0.0654	0.0977	0.2117	0.0307
EMPLOYEES_ON_NONFARM_04	0.7194	0.0057	0.0878	0.0005	0.0023	0.1938	-0.0245	-0.1156	-0.0096	-0.0239
EMPLOYEES_ON_NONFARM_02	0.1138	0.0022	0.0964	0.0163	0.0010	0.0771	-0.0151	-0.1211	0.0563	-0.0158
EMPLOYEES_ON_NONFARM_06	0.4423	0.0001	0.0521	0.0072	0.0085	0.1519	-0.0033	-0.0890	0.0375	-0.0458
EMPLOYEES_ON_NONFARM_05	0.6979	0.0082	0.0866	0.0031	0.0010	0.1908	-0.0295	-0.1148	-0.0244	-0.0157
EMPLOYEES_ON_NONFARM_07	0.5097	0.0041	0.0347	0.0411	0.0482	0.1631	0.0209	-0.0727	0.0893	-0.1093
EMPLOYEES_ON_NONFARM_08	0.5221	0.0097	0.0608	0.0960	0.0321	0.1651	0.0321	-0.0962	0.1366	-0.0893
EMPLOYEES_ON_NONFARM_09	0.5509	0.0059	0.1000	0.0475	0.0186	0.1696	0.0250	-0.1234	0.0960	-0.0678
EMPLOYEES_ON_NONFARM_10	0.3949	0.0144	0.0294	0.1160	0.0194	0.1436	0.0390	-0.0669	0.1501	-0.0694
EMPLOYEES_ON_NONFARM_12	0.0001	0.0071	0.0001	0.0011	0.0235	0.0025	-0.0273	0.0028	-0.0146	-0.0763
EMPLOYEES_ON_NONFARM_11	0.2352	0.0008	0.0146	0.0316	0.0237	0.1108	-0.0090	-0.0471	0.0783	-0.0766
INDEX_OF_AGGREGATE_WEEKL	0.4855	0.0218	0.0128	0.0653	0.0000	0.1592	0.0480	0.0441	0.1126	-0.0006
AVG_WEEKLY_HRS_OF_PROD_O	0.1049	0.0237	0.1506	0.0294	0.0158	0.0740	0.0501	0.1514	0.0756	0.0626
AVG_WEEKLY_HRS_OF_PRO01	0.1059	0.0088	0.1060	0.0032	0.0243	0.0743	0.0305	0.1270	0.0248	0.0777
AVERAGE_WEEKLY_HOURS__MF	0.1298	0.0086	0.1469	0.0001	0.0170	0.0823	0.0301	0.1495	-0.0045	0.0648
NAPM_EMPLOYMENT_INDEX	0.4923	0.0006	0.1640	0.0071	0.0000	0.1603	0.0079	-0.1580	0.0372	0.0007

Table 1 (continued)

Factor Loadings and their Interpretation

Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
	Univariate Regression R-square					Factor Loadings				
HOUSING_STARTS_NONFARM_1	0.0317	0.0192	0.1589	0.2282	0.0565	0.0407	0.0451	0.1555	0.2105	0.1184
HOUSING_STARTS_NORTHEAST	0.0102	0.0195	0.0588	0.0529	0.0191	0.0231	0.0454	0.0946	0.1014	0.0689
HOUSING_STARTS_MIDWEST	0.0042	0.0000	0.0625	0.1469	0.0218	0.0148	-0.0013	0.0975	0.1689	0.0736
HOUSING_STARTS_SOUTH	0.0237	0.0052	0.1119	0.1264	0.0136	0.0352	0.0234	0.1305	0.1567	0.0580
HOUSING_STARTS_WEST	0.0076	0.0076	0.0076	0.0076	0.0076	0.0199	0.0425	0.0299	0.0495	0.0887
HOUSING_AUTHORIZED__TOTA	0.0367	0.0942	0.1791	0.1300	0.0110	0.0438	0.0998	0.1651	0.1589	0.0522
HOUSES_AUTHORIZED_BY_BUI	0.0079	0.0109	0.0821	0.0590	0.0005	0.0203	0.0339	0.1118	0.1071	0.0111
HOUSES_AUTHORIZED_BY_01	0.0020	0.0364	0.0775	0.3124	0.0315	0.0103	0.0620	0.1086	0.2463	0.0884
HOUSES_AUTHORIZED_BY_02	0.0361	0.1145	0.1354	0.0605	0.0045	0.0434	0.1101	0.1435	0.1084	0.0335
HOUSES_AUTHORIZED_BY_03	0.0118	0.0001	0.0101	0.0023	0.0000	0.0248	0.0027	0.0391	-0.0211	-0.0016
PURCHASING_MANAGERS_IND	0.6090	0.0027	0.0757	0.0031	0.0035	0.1783	0.0170	-0.1073	0.0246	0.0293
NAPM_NEW_ORDERS_INDEX	0.5129	0.0061	0.0081	0.0000	0.0200	0.1636	0.0255	-0.0351	-0.0018	0.0704
NAPM_VENDOR_DELIVERIES_I	0.2421	0.0008	0.0657	0.0053	0.0022	0.1124	0.0095	-0.1000	0.0322	0.0236
NAPM_INVENTORIES_INDEX	0.2462	0.0000	0.1972	0.0304	0.0020	0.1134	-0.0009	-0.1732	0.0769	-0.0223
MFRS__NEW_ORDERS__CONSUM	0.1490	0.0014	0.1052	0.0311	0.0000	0.0882	0.0121	0.1265	-0.0778	0.0034
MFRS__NEW_ORDERS__DURABL	0.1024	0.0102	0.0362	0.0302	0.0039	0.0731	-0.0328	0.0742	-0.0766	-0.0311
MFRS__NEW_ORDERS__NONDEF	0.0494	0.0203	0.0041	0.0099	0.0000	0.0508	-0.0464	0.0249	-0.0437	0.0000
MFRS__UNFILLED_ORDERS__D	0.1077	0.0035	0.0913	0.0015	0.0451	0.0750	-0.0192	-0.1179	0.0171	-0.1057
MANUFACTURING_AND_TRADE_	0.2916	0.0010	0.2046	0.0023	0.0078	0.1234	-0.0105	-0.1765	0.0210	-0.0440
U__OF_MICH__INDEX_OF_CON	0.0001	0.0071	0.0058	0.0118	0.0120	-0.0018	-0.0274	0.0297	0.0478	-0.0546
RATIO__MFG__AND_TRADE_IN	0.0524	0.0006	0.4360	0.0000	0.0171	-0.0523	-0.0078	-0.2576	-0.0014	0.0650
MONEY_STOCK__M1	0.0020	0.0000	0.0848	0.0062	0.3042	-0.0102	0.0015	-0.1136	-0.0348	0.2746
MONEY_STOCK__M2	0.0000	0.0000	0.0954	0.0001	0.2843	-0.0011	-0.0001	-0.1205	0.0036	0.2655
MONEY_STOCK__CURRENCY_HE	0.0000	0.0011	0.0166	0.0069	0.0050	-0.0015	0.0106	-0.0503	-0.0366	0.0352
MONEY_SUPPLY__REAL_M2__A	0.0954	0.0447	0.0226	0.0001	0.0618	-0.0706	-0.0688	-0.0586	0.0042	0.1238
MONETARY_BASE__ADJ_FOR_R	0.0031	0.1496	0.0517	0.0225	0.0726	0.0128	-0.1258	-0.0887	-0.0661	0.1341
DEPOSITORY_INST_RESERVES	0.0016	0.0110	0.0938	0.0000	0.2846	0.0092	-0.0340	-0.1195	0.0017	0.2656
DEPOSITORY_INST_RESER01	0.0019	0.0724	0.0307	0.0489	0.0017	-0.0100	-0.0875	-0.0684	-0.0975	-0.0206
COMMERCIAL_AND_INDUSTRIA	0.0079	0.0112	0.0086	0.0084	0.1027	0.0203	0.0334	-0.0361	0.0405	0.1596
CHANGE_IN_COMMERCIAL_AND	0.0541	0.0122	0.2296	0.0308	0.0047	0.0531	-0.0359	-0.1869	0.0773	-0.0341
CONSUMER_CREDIT_OUTSTAND	0.0000	0.8631	0.0061	0.0176	0.0117	0.0003	0.0344	0.0306	0.0585	-0.0538
RATIO__CONSUMER_INSTALLM	0.0427	0.0025	0.0150	0.0023	0.0405	-0.0472	0.0162	-0.0477	-0.0213	-0.1002
S_P_DIV_YIELD	0.0078	0.0482	0.0791	0.0150	0.0269	-0.0202	-0.0714	-0.1097	-0.0540	0.0816
S_P_PE_RATIO	0.0831	0.0155	0.1557	0.0008	0.0094	-0.0659	0.0404	0.1539	-0.0128	-0.0482
NOMINAL_E_ECTIVE_EXCHANG	0.0000	0.0630	0.0047	0.0185	0.0256	0.0014	-0.0816	0.0266	0.0599	-0.0797
FOREIGN_EXCHANGE_RATE__S	0.0010	0.0746	0.0079	0.0002	0.0640	0.0074	-0.0888	0.0346	-0.0061	-0.1259
FOREIGN_EXCHANGE_RATE__J	0.0051	0.0133	0.0125	0.0004	0.0411	0.0163	-0.0374	0.0436	0.0088	-0.1010
FOREIGN_EXCHANGE_RATE__U	0.0108	0.0694	0.0000	0.0146	0.0476	0.0237	0.0857	0.0013	0.0533	0.1086
FOREIGN_EXCHANGE_RATE__C	0.0010	0.1147	0.0001	0.0000	0.0015	-0.0074	-0.1101	-0.0035	0.0006	0.0192
PRODUCER_PRICE_INDEX__FI	0.0024	0.6113	0.0141	0.0185	0.0000	-0.0111	0.2543	-0.0464	-0.0600	-0.0012
PRODUCER_PRICE_INDEX_01	0.0030	0.6158	0.0137	0.0188	0.0005	-0.0125	0.2552	-0.0456	-0.0605	-0.0115
PRODUCER_PRICE_INDEX_I_N	0.0008	0.6118	0.0088	0.0362	0.0001	-0.0064	0.2544	-0.0366	-0.0838	-0.0057
PRODUCER_PRICE_INDEX__CR	0.0015	0.2532	0.0158	0.0740	0.0057	-0.0088	0.1636	-0.0490	-0.1199	-0.0376
CPI_U__ALL_ITEMS	0.0039	0.8440	0.0119	0.0032	0.0018	-0.0143	0.3022	-0.0425	-0.0249	0.0209
CPI_U__APPAREL__UPKEEP	0.0004	0.0040	0.0058	0.0087	0.0220	0.0046	0.0206	0.0297	-0.0411	-0.0739
CPI_U__TRANSPORTATION	0.0049	0.7877	0.0036	0.0030	0.0046	-0.0160	0.2887	-0.0235	-0.0240	0.0336
CPI_U__MEDICAL_CARE	0.0021	0.0000	0.0009	0.0050	0.0351	-0.0104	0.0005	-0.0114	0.0313	-0.0933
CPI_U__DURABLES	0.0089	0.0215	0.0072	0.0193	0.0063	-0.0215	0.0477	-0.0332	-0.0613	-0.0395
CPI_U__SERVICES	0.0001	0.0721	0.0046	0.0061	0.0278	0.0027	0.0873	-0.0266	-0.0345	-0.0830
CPI_U__ALL_ITEMS_LESS_FO	0.0039	0.8645	0.0085	0.0091	0.0005	-0.0142	0.2988	-0.0359	-0.0420	0.0110
CPI_U__ALL_ITEMS_LESS_SH	0.0026	0.8671	0.0052	0.0053	0.0047	-0.0116	0.3029	-0.0283	-0.0321	0.0343
CPI_U__ALL_ITEMS_LESS_MI	0.0028	0.8645	0.0101	0.0057	0.0027	-0.0120	0.3024	-0.0392	-0.0333	0.0259
PCE__IMPL_PR_DEFLATOR_PC	0.0030	0.7225	0.0000	0.0122	0.0443	-0.0125	0.2765	0.0006	-0.0486	-0.1048
PCE__IMPL_PR_DEFLATOR01	0.0017	0.0495	0.0001	0.0015	0.0734	-0.0095	0.0724	0.0037	-0.0172	-0.1348
PCE__IMPL_PR_DEFLATOR02	0.0028	0.7920	0.0053	0.0024	0.0105	-0.0120	0.2894	-0.0285	-0.0217	0.0510
PCE__IMPL_PR_DEFLATOR03	0.0001	0.0177	0.0141	0.0171	0.2442	-0.0019	0.0433	0.0463	-0.0576	-0.2460
AVG_HOURLY_EARNINGS_OF_P	0.0005	0.0006	0.0088	0.0187	0.0409	0.0053	-0.0078	-0.0366	-0.0602	0.1007
AVG_HOURLY_EARNINGS_001	0.0049	0.0022	0.0461	0.0178	0.0006	-0.0160	-0.0151	-0.0838	-0.0587	-0.0126
AVG_HOURLY_EARNINGS_002	0.0051	0.0059	0.0002	0.0372	0.0794	0.0162	-0.0251	0.0048	-0.0850	0.1403
GSERFCI_INDEX	0.2474	0.0035	0.0671	0.0266	0.0031	-0.1136	-0.0194	0.1011	-0.0719	0.0277
HISTORICAL_UNCERTAINTY	0.0640	0.0005	0.0031	0.0085	0.0343	-0.0578	-0.0075	-0.0218	-0.0407	0.0922
LIQUIDITY_RISK_FACTOR	0.0273	0.0094	0.0235	0.0015	0.0727	0.0377	-0.0315	0.0598	-0.0169	-0.1343

Table 2

Summary Statistics (January 1989-December 2011)

The following tables refer to all the series used in the parts of the paper based on the sample January 31, 1989 - December 30, 2011. Panel A shows mean, median, standard deviation, skewness, excess kurtosis, and the p-value of the Jarque-Bera (JB) test. The JB statistic is used to test the hypothesis of normality of the series, hence a p-value lower than α% implies the rejection of the null hypothesis at α % confidence level. Panel B shows the correlations among the series.

	Mean	**Median**	**Std. Dev.**	**Skewness**	**Excess Kurtosis**	**JB p-value**
Factor 1	0.0000	0.0080	0.4385	-1.6940	4.8482	0.0000
Factor 2	0.0000	0.0014	0.3080	-0.7455	4.2518	0.0000
Factor 3	0.0000	-0.0001	0.2568	0.6517	2.5557	0.0000
HP Factor	0.0036	0.0049	0.0442	-0.1246	1.0115	0.0019
Basis Factor	0.0098	0.0109	0.0448	-0.0093	1.3934	0.0000
Momentum Factor	0.0088	0.0091	0.0503	0.3528	3.3293	0.0000
Agricolture and Livestock	-0.0002	0.0003	0.0428	-0.0876	1.5607	0.0000
Precious Metals	0.0030	-0.0024	0.0469	0.0980	1.2784	0.0001
Industrial Metals	0.0008	-0.0007	0.0606	-0.1039	1.4521	0.0000
Energy	0.0071	0.0086	0.0904	0.3851	1.5595	0.0000
10Y Treasury Bonds	0.0023	0.0020	0.0019	0.0016	-0.7515	0.0389
Aaa Corporate Bonds	0.0033	0.0031	0.0019	-0.0005	-0.7615	0.0356
Baa Corporate Bonds	0.0042	0.0039	0.0019	0.0104	-0.7375	0.0437
Value-Weighted Equity CRSP	0.0054	0.0110	0.0447	-0.6165	1.0217	0.0000

Panel B

	Factor 1	Factor 2	Factor 3	HP	Basis	Momentum	AL	PM	IN	EN	Treasuries	Aaa Bonds	Baa Bonds	VW CRSP
Factor 1	1.00	0.00	0.00	-0.02	-0.11	-0.08	0.08	-0.05	0.19	0.10	-0.08	-0.08	-0.07	0.09
Factor 2		1.00	0.00	0.07	0.05	0.01	0.08	0.24	0.16	0.27	0.03	0.04	0.06	0.04
Factor 3			1.00	0.03	0.03	-0.02	0.00	-0.09	0.14	-0.04	0.12	0.13	0.15	0.10
HP Factor				1.00	0.03	0.42	0.15	0.24	0.18	0.09	0.13	0.13	0.14	0.02
Basis Factor					1.00	0.14	0.06	-0.09	-0.01	0.07	-0.05	-0.04	-0.04	0.02
Momentum Factor						1.00	-0.07	0.11	0.06	0.06	-0.10	-0.10	-0.10	-0.02
Agricolture and Livestock							1.00	0.29	0.31	0.16	0.12	0.12	0.13	0.22
Precious Metals								1.00	0.26	0.21	0.21	0.21	0.21	0.03
Industrial Metals									1.00	0.29	0.09	0.10	0.12	0.38
Energy										1.00	-0.01	0.00	0.01	0.09
10Y Treasury Bonds											1.00	1.00	1.00	-0.02
Aaa Corporate Bonds												1.00	1.00	-0.01
Baa Corporate Bonds													1.00	0.00
Value-weighted Equity CRSP														1.00

Table 3

Model Selection

The table shows the statistics used to select the best fitting model. SIC is the Bayes-Schwarz information criterion, AIC is the Akaike information criterion, and HQIC is the Hannan-Quinn information criterion.

	VAR order	Maximized Log-Likelihood	Average Log-Likelihood	No. Parameters	Saturation ratio	SIC	AIC	HQIC
	Single-State Models							
Constant	0	9095.61	3.67	1	2475	-7.347	-7.349	-7.348
	Three-factor models							
3 factors, single state, full VAR model	1	8056.37	2.66	40	75.63	-5.221	-5.300	-5.271
3 factors, single state, block Factor VAR	1	7964.23	2.63	16	189.06	-5.223	-5.255	-5.244
3 factors,single state, diagonal Factor VAR	1	7828.44	2.59	10	302.50	-5.149	-5.169	-5.162
	Five-factor models							
5 factors, single state, full VAR model	1	7537.22	2.11	76	47.04	-4.043	-4.174	-4.127
5 factors, single state, block Factor VAR	1	7432.32	2.08	36	99.31	-4.076	-4.138	-4.116
5 factors, single state, diagonal Factor VAR	1	7187.82	2.01	16	223.44	-3.985	-4.012	-4.002
	Commodity-based models (three factors)							
Only commodity factors, all assets, single state, full VAR model	1	11634.17	3.85	40	75.63	-7.586	-7.666	-7.637
Only commodity factors, all assets, single state, block Factor VAR	1	11283.27	3.73	16	189.06	-7.418	-7.449	-7.438
Only commodity factors, all assets, single state, diagonal Factor VAR	1	11226.34	3.71	10	302.50	**-7.767**	-7.416	-7.409
	All-factor models (eight factors)							
All factors, all assets, single state, full VAR model	1	10340.68	2.69	97	39.69	-5.164	-5.321	-5.265
All factors, all assets, single state, block Factor VAR	1	10165.89	2.64	49	78.57	-5.176	-5.256	-5.227
All factors, all assets, single state, diagonal VAR	1	10006.95	2.60	19	202.63	-5.158	-5.189	-5.178

Table 3 (continued)

Model Selection

	VAR order	Maximized Log-Likelihood	Average Log-Likelihood	No. Parameters	Saturation ratio	SIC	AIC	HQIC
Regime Switching Models (Two States)								
	Three-factor models							
3 factors, 2 states, full VAR model	1	8228.45	2.72	82	36.89	-5.223	-5.386	-5.327
3 factors, 2 states, block Factor VAR	1	8218.11	2.72	34	88.97	-5.343	-5.411	-5.387
3 factors, 2 states, diagonal Factor VAR	1	8080.99	2.67	22	137.50	-5.285	-5.328	-5.313
	Five-factor models							
5 factors, 2 states, full VAR model	1	7866.04	2.20	154	23.21	-4.048	-4.314	-4.219
5 factors, 2 states, block Factor VAR	1	7720.26	2.16	74	48.31	-4.150	-4.278	-4.232
5 factors, 2 states, diagonal Factor VAR	1	7513.10	2.10	34	105.15	-4.125	-4.184	-4.163
	Commodity-based models (three factors)							
Only commodity factors, all assets, 2 states, full VAR model	1	11795.67	3.90	82	36.89	-7.582	**-7.745**	**-7.686**
Only commodity factors, all assets, 2 states, block Factor VAR	1	11551.90	3.82	34	88.97	-7.548	-7.615	-7.591
Only commodity factors, all assets, 2 states,diagonal Factor VAR	1	11548.22	3.82	22	137.50	-7.577	-7.621	-7.605
	All-factor models (eight factors)							
All factors, all assets, 2 states, full VAR model	1	10566.13	2.74	196	19.64	-5.069	-5.387	-5.274
All factors, all assets, 2 states, block Factor VAR	1	10431.30	2.71	100	38.50	-5.204	-5.367	-5.309
All factors, all assets, 2 states, diagonal Factor VAR	1	10306.39	2.68	40	96.25	-5.268	-5.333	-5.310
Regime Switching Models (Three States)								
	Three-factor models							
3 factors, 3 states, full VAR model	1	8344.14	2.76	126	24.01	-5.183	-5.433	-5.343
3 factors, 3 states, block Factor VAR	1	8301.35	2.74	54	56.02	-5.345	-5.453	-5.414
3 factors, 3 states, diagonal Factor VAR	1	8182.93	2.71	36	84.03	-5.315	-5.386	-5.361
	Five-factor models							
5 factors, 3 states, full VAR model	1	7978.98	2.23	234	15.28	-3.928	-4.333	-4.189
5 factors, 3 states, block Factor VAR	1	7821.19	2.19	114	31.36	-4.115	-4.312	-4.241
5 factors, 3 states, diagonal Factor VAR	1	7609.888	2.13	54	66.20	-4.134	-4.227	-4.194
	Commodity-based models (three factors)							
Only commodity factors, all assets, 3 states, full VAR model	1	11795.26	3.90	126	24.01	-7.465	-7.715	-7.625
Only commodity factors, all assets, 3 states, block Factor VAR	1	11658.87	3.85	54	56.02	-7.565	-7.673	-7.634
Only commodity factors, all assets, 3 states,diagonal Factor VAR	1	11697.49	3.87	36	84.03	-7.638	-7.710	-7.684
	All-factor models (eight factors)							
All factors, all assets, 3 states, full VAR model	1	10741.35	2.79	297	12.96	-4.943	-5.426	-5.254
All factors, all assets, 3 states, block Factor VAR	1	10603.56	2.75	153	25.16	-5.180	-5.429	-5.341
All factors, all assets, 3 states, diagonal Factor VAR	1	10475.56	2.72	63	61.11	-5.307	-5.409	-5.373

Table 4

Model Estimates- Only Commodity Factors, Single State, Diagonal VAR

	Regime 1				Regime 2				Δ Coeff.
	Coefficient	Std. Error	t-Statistic	p-value	Coefficient	Std. Error	t-Statistic	p-value	
Factor VAR Coefficients									
HP: AR(1) Coeff.	0.0337	0.0219	1.5354	0.1248	—	—	—	—	—
HP: AR(1) Intercept	0.0033	0.0025	1.2985	0.1942	—	—	—	—	—
Basis : AR(1) Coeff.	-0.0221	0.0565	-0.3907	0.6960	—	—	—	—	—
Basis: AR(1) Intercept	0.0097	0.0027	3.5609	0.0004	—	—	—	—	—
Momentum: AR(1) Coeff.	0.0869	0.0527	1.6493	0.0992	—	—	—	—	—
Momentum: AR(1) Intercept	0.0081	0.0030	2.6773	0.0075	—	—	—	—	—
SDF: Loading on HP	0.1156	0.0028	40.8860	0.0000	—	—	—	—	—
SDF: Loading on Basis	0.0119	0.0026	4.6668	0.0000	—	—	—	—	—
SDF: Loading on Momentum	-0.0178	0.0025	-7.0921	0.0000	—	—	—	—	—
SDF: Intercept	0.0018	0.0003	6.1760	0.0000	—	—	—	—	—

Residual Covariance Matrix of factors and test assets

	HP	Basis	Momentum	Agr. & Livestock	Precious	Industrials	Energy	10Y Treasuries	Aaa Corporate	Baa Corporate	VW CRSP
HP	0.0020	0.0001	0.0010	0.0001	0.0003	0.0003	0.0001	-0.0002	-0.0002	-0.0002	-0.0002
Basis		0.0020	0.0003	0.0001	-0.0002	-0.0001	0.0003	0.0000	0.0000	0.0000	0.0000
Momentum			0.0025	-0.0002	0.0002	0.0001	0.0002	-0.0001	-0.0001	-0.0001	-0.0001
Agricolture & Livestock				0.0018	0.0005	0.0007	0.0005	0.0000	0.0000	0.0000	0.0004
Precious Metals					0.0021	0.0006	0.0008	0.0000	0.0000	0.0000	0.0000
Industrial Metals						0.0036	0.0015	0.0000	0.0000	0.0000	0.0010
Energy							0.0081	0.0000	0.0000	0.0000	0.0003
10Y Treasury Bonds								0.0000	0.0000	0.0000	0.0000
Aaa Corporate Bonds									0.0000	0.0000	0.0000
Baa Corporate Bonds										0.0000	0.0000
VW Equity CRSP											0.0020

Table 5

Model Estimates- Only commodity factors, Full VAR, Two regimes

	Regime 1						Regime 2				Δ Coeff.
	Coefficient	Std. Error	t-Statistic	p-value			Coefficient	Std. Error	t-Statistic	p-value	
	Factor VAR Coefficients and Estimated Transition Matrix										
[HP/HP] Coeff.	0.204	0.041	4.952	0.000			-0.049	0.031	-1.546	0.122	0.253
[HP/Basis] Coeff.	0.051	0.035	1.467	0.143			-0.002	0.028	-0.085	0.932	0.054
[HP/Momentum] Coeff.	-0.031	0.038	-0.810	0.418			0.001	0.027	0.036	0.971	-0.032
HP: Intercept	-0.186	0.029	-6.399	0.000			0.002	0.016	0.098	0.922	-0.187
[Basis/HP] Coeff.	0.256	0.048	5.325	0.000	**Transition Matrix**		-0.038	0.067	-0.565	0.572	0.294
[Basis/Basis] Coeff.	0.012	0.040	0.291	0.771	0.806	0.089	0.035	0.059	0.584	0.559	-0.023
[Basis/Momentum] Coeff.	-0.101	0.044	-2.291	0.022	0.194	0.911	0.009	0.058	0.150	0.881	-0.110
Basis: intercept	-0.215	0.050	-4.329	0.000	**Implied Durations**		0.208	0.066	3.136	0.002	-0.423
[Momentum/HP] Coeff.	-0.024	0.034	-0.723	0.470	Regime 1:	5.16	0.048	0.033	1.465	0.143	-0.072
[Momentum/Basis] Coeff.	-0.225	0.044	-5.156	0.000	Regime 2:	11.22	-0.133	0.078	-1.710	0.087	-0.092
[Momentum/Momentum] Coef	0.015	0.036	0.420	0.674	**Ergodic Probs.:**		0.002	0.069	0.033	0.973	0.013
Momentum Intercept	-0.027	0.037	-0.738	0.460	Regime 1:	0.315	-0.070	0.077	-0.913	0.361	0.043
SDF: Loading on HP	0.014	0.030	0.458	0.647	Regime 2:	0.685	0.045	0.038	1.193	0.233	-0.031
SDF: Loading on Basis	-0.015	0.006	-2.383	0.017			0.154	0.002	68.526	0.000	-0.169
SDF: Loading on Momentum	0.102	0.004	24.866	0.000			-0.015	0.002	-7.518	0.000	0.118
SDF: Intercept	0.117	0.005	25.618	0.000			-0.051	0.002	-26.268	0.000	0.168

Residual Covariance Matrix of factors and test assets (Regime 1 above; regime 2 below)

	HP	Basis	Momentum	Agr. & Livestock	Precious	Industrials	Energy	10Y Treasuries	Aaa Corporate	Baa Corporate	VW CRSP
HP	0.00048|0.00229	-0.0001	0.0003	-0.0001	-0.0001	-0.0002	0.0004	0.0000	0.0000	0.0000	-0.0001
Basis	0.0001	0.00119|0.00211	0.0002	0.0002	-0.0003	-0.0001	0.0003	-0.0001	-0.0001	-0.0001	-0.0001
Mometum	0.0012	0.0003	0.00105|0.002953	-0.0003	-0.0002	0.0001	0.0010	-0.0001	-0.0001	-0.0001	-0.0001
Agricolture & Livestock	0.0001	0.0000	-0.0001	0.00086|0.00219	0.0001	-0.0001	-0.0003	0.0000	0.0000	0.0000	-0.0001
Precious Metals	0.0004	-0.0002	0.0005	0.0007	0.00079|0.00270	0.0002	0.0000	0.0000	0.0000	0.0000	-0.0001
Industrial Metals	0.0005	-0.0002	0.0002	0.0011	0.0008	0.00269|0.00390	0.0005	0.0000	0.0000	0.0000	0.0002
Energy	0.0000	0.0001	-0.0001	0.0009	0.0011	0.0019	0.00492|0.009513	-0.0001	-0.0001	-0.0001	-0.0002
10Y Treasury Bonds	-0.0003	0.0000	0.0000	0.0000	0.0000	-0.0001	0.0000	0.000037|0.000045	0.0000	0.0000	0.0000
Aaa Corporate Bonds	-0.0003	0.0000	0.0000	0.0000	0.0000	-0.0001	0.0000	0.0000	0.00004|0.000045	0.0000	0.0000
Baa Corporate Bonds	-0.0003	0.0000	0.0000	0.0000	0.0000	-0.0001	0.0000	0.0000	0.0000	0.000046|0.000047	0.0000
VW Equity CRSP	-0.0002	0.0001	-0.0002	0.0006	0.0001	0.0013	0.0006	0.0000	0.0000	0.0000	0.00149|0.00224

Table 6

Observed and Implied Correlations- Comparing Different Models

The tables reports observed and model-implied correlations. The check column indicates whether the model correlation falls inside the 90% confidence interval around the observed value.

			Empirical	90% CI - lower bound	90% CI - upper bound	Only commodity factors, diagonal VAR, single state	3 macro factors, diagonal VAR, single state	All factors, diagonal VAR, single state	Only commodity factors, Full VAR, two regimes
Mean	Factor 1		0.000	-0.436	0.436		☒	☑	
	Factor 2		0.000	-0.306	0.306		☒	☑	
	Factor 3		0.000	-0.255	0.255		☒	☑	
	HP		0.004	-0.001	0.008	☑		☑	☒
	Basis		0.010	0.005	0.014	☑		☑	☑
	Momentum		0.009	0.004	0.014	☑		☒	☒
	Agri&Livestock		0.000	-0.004	0.004	☑	☒	☑	☑
	Precious Metals		0.003	-0.002	0.008	☑	☒	☑	☑
	Industrial Metals		0.001	-0.005	0.007	☑	☒	☑	☑
	Energy		0.007	-0.002	0.016	☑	☑	☑	☑
	10Y Treasuries		0.002	0.002	0.002	☑	☒	☒	☒
	Aaa Corporate Bonds		0.003	0.003	0.004	☒	☒	☒	☒
	Baa Corporate Bonds		0.004	0.004	0.004	☒	☒	☒	☒
	CRSP Value-weighted Equity Ptf.		0.005	0.001	0.010	☑	☑	☑	☑
Standard Deviation	Factor 1		4.377	4.307	4.448		☒	☒	
	Factor 2		3.075	3.004	3.145		☒	☑	
	Factor 3		2.563	2.493	2.634		☒	☒	
	HP		0.044	-0.026	0.115	☑		☑	☒
	Basis		0.045	-0.026	0.115	☑		☑	☑
	Momentum		0.050	-0.020	0.121	☑		☒	☑
	Agri&Livestock		0.043	-0.028	0.113	☑	☑	☑	☑
	Precious Metals		0.047	-0.024	0.117	☑	☑	☑	☑
	Industrial Metals		0.060	-0.010	0.131	☑	☑	☑	☑
	Energy		0.090	0.020	0.161	☑	☑	☑	☑
	10Y Treasuries		0.002	-0.069	0.072	☑	☑	☑	☑
	Aaa Corporate Bonds		0.002	-0.069	0.072	☑	☑	☑	☑
	Baa Corporate Bonds		0.002	-0.069	0.072	☑	☑	☑	☑
	CRSP Value-weighted Equity Ptf.		0.045	-0.026	0.115	☑	☑	☑	☑
Skewness	Factor 1		-1.703	-1.947	-1.460		☒	☒	
	Factor 2		-0.750	-0.993	-0.506		☒	☒	
	Factor 3		0.655	0.412	0.899		☒	☒	
	HP		-0.125	-0.369	0.118	☑		☑	☒
	Basis		-0.009	-0.253	0.234	☑		☑	☒
	Momentum		0.355	0.111	0.598	☒		☒	☒
	Agri&Livestock		-0.088	-0.332	0.156	☑	☒	☑	☑
	Precious Metals		0.099	-0.145	0.342	☑	☒	☑	☑
	Industrial Metals		-0.104	-0.348	0.139	☑	☒	☑	☑
	Energy		0.387	0.144	0.631	☒	☒	☒	☒
	10Y Treasuries		0.002	-0.242	0.245	☑	☒	☑	☒
	Aaa Corporate Bonds		-0.001	-0.244	0.243	☑	☒	☑	☒
	Baa Corporate Bonds		0.010	-0.233	0.254	☑	☒	☑	☒
	CRSP Value-weighted Equity Ptf.		-0.620	-0.864	-0.376	☒	☒	☒	☒
Excess Kurtosis	Factor 1		4.959	4.472	5.447		☒	☒	
	Factor 2		4.352	3.864	4.839		☒	☒	
	Factor 3		2.625	2.137	3.112		☒	☒	
	HP		1.052	0.565	1.540	☒		☒	☒
	Basis		1.441	0.954	1.928	☒		☒	☒
	Momentum		3.412	2.925	3.900	☒		☒	☒
	Agri&Livestock		1.611	1.124	2.099	☒	☑	☒	☒
	Precious Metals		1.324	0.836	1.811	☒	☒	☒	☒
	Industrial Metals		1.501	1.013	1.988	☒	☑	☒	☒
	Energy		1.610	1.123	2.098	☒	☒	☒	☒
	10Y Treasuries		-0.743	-1.231	-0.256	☒	☒	☒	☒
	Aaa Corporate Bonds		-0.753	-1.241	-0.266	☒	☒	☒	☒
	Baa Corporate Bonds		-0.729	-1.216	-0.241	☒	☒	☒	☒
	CRSP Value-weighted Equity Ptf.		1.063	0.575	1.550	☒	☒	☒	☒
	Agri&Livestock	Precious Metals	0.292	0.193	0.392	☑	☒	☑	☑
		Industrial Metals	0.305	0.206	0.405	☑	☑	☑	☑
		Energy	0.159	0.060	0.259	☑	☑	☑	☑
		10Y Treasuries	0.121	0.022	0.221	☑	☒	☑	☑
		Aaa Corp. Bonds	0.123	0.024	0.223	☑	☒	☑	☑
		Baa Corp. Bonds	0.129	0.030	0.229	☑	☒	☒	☑
		CRSP Equities	0.221	0.122	0.321	☑	☒	☑	☑
	Precious Metals	Industrial Metals	0.257	0.157	0.356	☑	☒	☑	☑
		Energy	0.210	0.111	0.310	☑	☑	☑	☑
		10Y Treasuries	0.209	0.109	0.308	☑	☒	☑	☒
		Aaa Corp. Bonds	0.209	0.110	0.309	☑	☒	☑	☒
		Baa Corp. Bonds	0.213	0.113	0.312	☑	☒	☑	☒
		CRSP Equities	0.026	-0.074	0.125	☑	☒	☑	☑
	Industrial Metals	Energy	0.287	0.188	0.387	☑	☑	☑	☑
		10Y Treasuries	0.088	-0.011	0.188	☑	☒	☑	☑
		Aaa Corp. Bonds	0.099	-0.001	0.198	☑	☒	☑	☑
		Baa Corp. Bonds	0.116	0.017	0.216	☑	☒	☑	☑
		CRSP Equities	0.381	0.281	0.480	☑	☑	☑	☑
	Energy	10Y Treasuries	-0.009	-0.108	0.091	☑	☒	☑	☑
		Aaa Corp. Bonds	0.001	-0.099	0.100	☑	☒	☑	☑
		Baa Corp. Bonds	0.014	-0.086	0.113	☑	☒	☑	☑
		CRSP Equities	0.090	-0.010	0.189	☑	☑	☑	☑
	10Y Treasuries	Aaa Corp. Bonds	0.999	0.899	1.098	☒	☑	☑	☑
		Baa Corp. Bonds	0.997	0.897	1.096	☒	☑	☒	☑
		CRSP Equities	-0.019	-0.119	0.080	☑	☒	☑	☒
	Aaa Bonds	Baa Corp. Bonds	0.999	0.899	1.098	☒	☑	☑	☑
		CRSP Equities	-0.010	-0.110	0.089	☑	☒	☑	☒
	Baa bonds	CRSP Equities	0.000	-0.100	0.099	☑	☒	☑	☒

Figure 1

Observed and Implied Correlations: Only Commodity Factors, Single State, Diagonal VAR

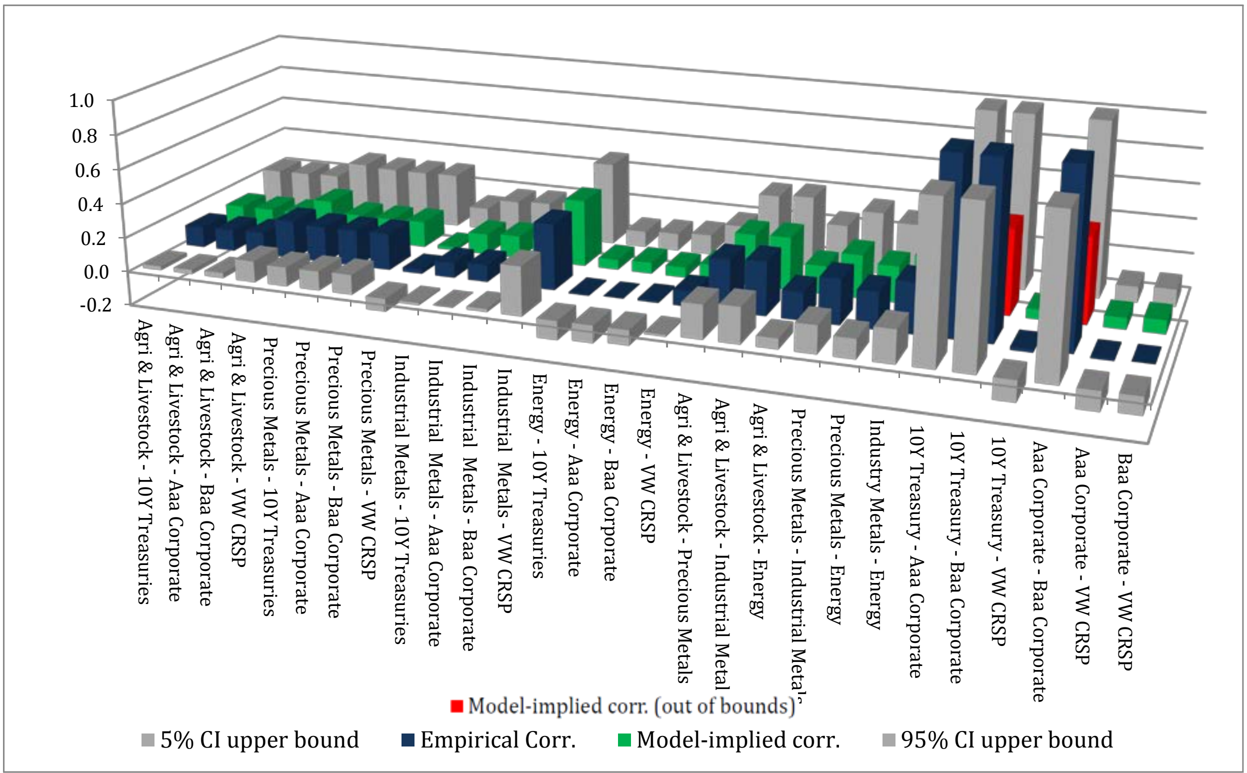


Figure 2

Observed and Implied Correlations: Three macro factors, Single State, Diagonal VAR

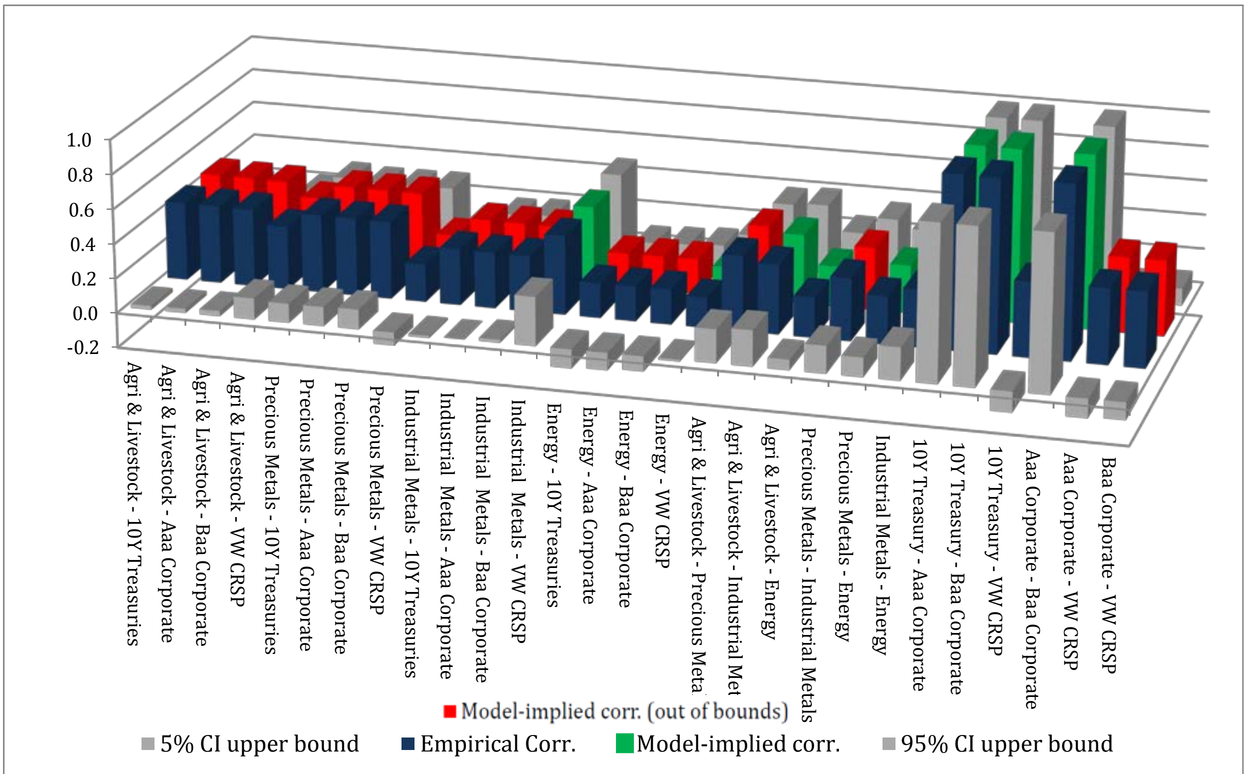


Figure 3

Observed and Implied Correlations: All factors (macro and commodity-specific), Single State, Diagonal VAR

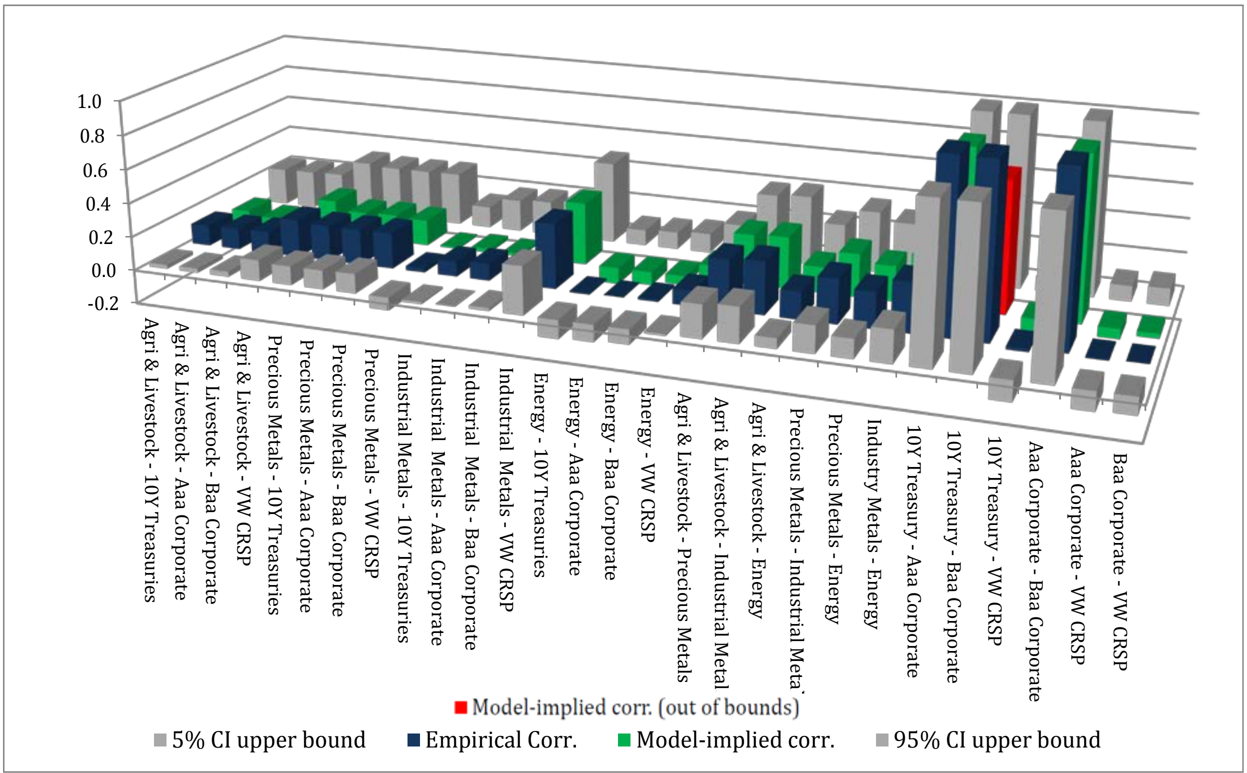
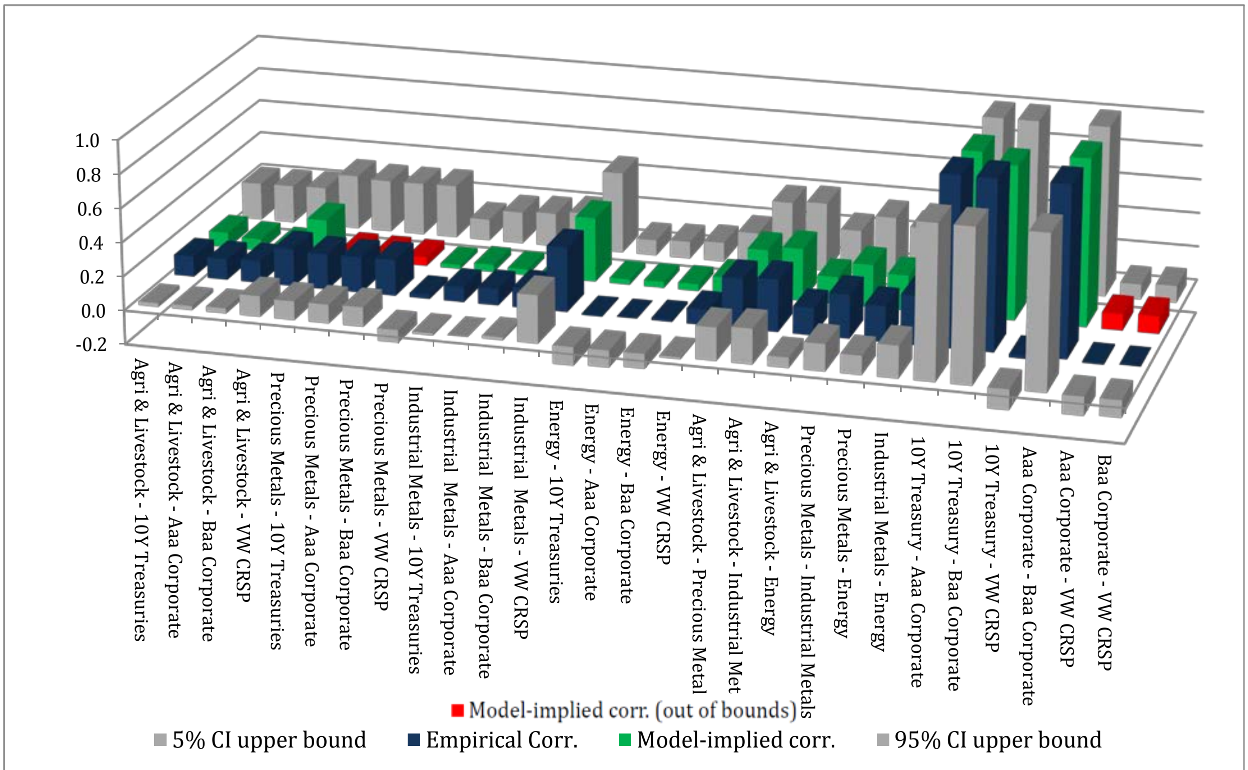


Figure 4

Observed and Implied Correlations: Only commodity factors, Full VAR, Two Regimes



Appendix A - Derivation of the Conditional Mean Models for Asset Returns

Given the general log-linear structure for the SDF in (3) and (9), we define $m_{t+1} \equiv lnM_{t+1}$ and $u_{i,t+1} \equiv lnM_{t+1}(1+R_{i,t+1}) = m_{t+1} + r_{i,t+1}$. In the following, also for identification purposes when the model is generalized to the HMM case, we assume that the K factors are observable. Positivity of M_{t+1} ensures the absence of arbitrage opportunities. If the random vector $\boldsymbol{y}_{t+1} \equiv [m_{t+1}\, f_{1,t+1}\, f_{2,t+1}\, \ldots\, f_{K,t+1}\;\; r_{1,t+1}\;\; r_{2,t+1}\; \ldots\; r_{n,t+1}]'$, with $\Psi_t \equiv (\{\boldsymbol{y}_{t-s}\}_{s\geq 0}, S_t)$ has a stationary multivariate Gaussian distribution with conditional mean $\boldsymbol{\mu}_{y,t}$ and (constant) covariance matrix $\boldsymbol{\Sigma}$, then $u_{i,t+1}$ will be a weighted sum of normal random variates and as such will also have a Gaussian distribution with mean $E[m_{t+1}|\Psi_t] + E[r_{i,t+1}|\Psi_t]$ and variance $\sigma_m^2 + \sigma_i^2 + 2\sigma_{i,m}$. Because $\Psi_t \subset \Im_t$, basic properties of conditional expectations yield

$$E[(U_{i,t+1}|\Im_t)|\Psi_t] = E[U_{i,t+1}|\Psi_t] = E[\exp(u_{i,t+1})|\Psi_t]$$
$$= exp\left\{E[m_{t+1}|\Psi_t] + E[r_{i,t+1}|\Psi_t] + \frac{1}{2}\sigma_m^2 + \frac{1}{2}\sigma_i^2 + \sigma_{i,m}\right\} = 1,$$

as $U_{i,t+1}$ has a log-normal distribution. Taking logs of both sides, reveals that

$$E[m_{t+1}|\Psi_t] + E[r_{i,t+1}|\Psi_t] = -\frac{1}{2}\sigma_m^2 - \frac{1}{2}\sigma_i^2 - \sigma_{i,m}$$

Therefore the random variable $v_{i,t+1} \equiv u_{i,t+1} - E[u_{i,t+1}|\Psi_t] = u_{i,t+1} - E[m_{t+1}|\Psi_t] - E[r_{i,t+1}|\Psi_t]$ is characterized as

$$v_{i,t+1} = m_{t+1} + r_{i,t+1} + \frac{1}{2}\sigma_m^2 + \frac{1}{2}\sigma_i^2 + \sigma_{i,m}$$

and $E[v_{i,t+1}|\Psi_t] = 0$ (this implies zero serial correlation) with

$$E[r_{i,t+1}|\Psi_t] = -E[m_{t+1}|\Psi_t] - \tfrac{1}{2}\sigma_m^2 - \tfrac{1}{2}\sigma_i^2 - \sigma_{i,m}.$$

At this point, recall that our expression for the SDF implies that

$$E[m_{t+1}|\Psi_t] = \gamma_0 + \sum_{j=1}^{K} \gamma_j f_{j,t+1}.$$

Therefore

$$m_{t+1} = \gamma_0 + \sum_{j=1}^{K} \gamma_j f_{j,t+1} + w_{t+1} \qquad (w_{t+1} \equiv m_{t+1} - E[m_{t+1}|\Psi_t])$$
$$f_{j,t+1} = \varphi_{j,0} + \sum_{k=1}^{K}\sum_{p=1}^{P} \varphi_{j,k,p} f_{k,t+1-p} + \delta_{j,t+1}, \qquad j = 1, \ldots, K \qquad (7)$$
$$r_{i,t+1} = \underbrace{(0.5\sigma_m^2 + 0.5\sigma_i^2 + \sigma_{i,m})}_{\mu_i} + m_{t+1} + v_{i,t+1} = \mu_i + m_{t+1} + v_{i,t+1}, \quad i = 1, \ldots, n$$

is our full multivariate model, spanning both the SDF linear function of observable factors and asset returns. As invoked in the main text, the extension of this framework to the HMM case appears to be straightforward.

Appendix B

List of (Macroeconomic) Variables Used to Extract Principal Components

Grouping	Variable	Acronym	Sample period	Notes
GROUP I Labor and Income	Personal Income	PI	1983:01-2011:12	Δln
	Personal Income Less Transfer Payments	PI_LESS_TRANSFERS	1983:01-2011:12	Δln
	Real Consumption	REAL_CONSUMPTION	1983:01-2011:12	Δln
	Manufacturing and Trade Sales	MANUFACTURIN__TRADE_SAL	1983:01-2011:12	Δln
	Sales of Retail Stores	RETAIL_SALES	1983:01-2011:12	Δln
	Industrial Production Index - Total Index	IP__TOTAL	1983:01-2011:12	Δln
	Industrial Production Index - Products, Total	IP__PRODUCTS	1983:01-2011:12	Δln
	Industrial Production Index - Final Products	IP__FINAL_PROD	1983:01-2011:12	Δln
	Industrial Production Index - Consumer Goods	IP__DURABLE_CONSUMER_GOO	1983:01-2011:12	Δln
	Industrial Production Index - Durable Consumer Goods	IP__DURABLE_CONSUMER_GOO	1983:01-2011:12	Δln
	Industrial Production Index - Nondurable Consumer Goods	IP__CONS_NONDBLE	1983:01-2011:12	Δln
	Industrial Production Index - Business Equipment	IP__BUS_EQPT	1983:01-2011:12	Δln
	Industrial Production Index - Materials	IP__MATERIALS	1983:01-2011:12	Δln
	Industrial Production Index - Durable Goods Materials	IP__DURABLE_GOODS_MATERI	1983:01-2011:12	Δln
	Industrial Production Index - Nondurable Goods Materials	IP__NONDURABLE_GOODS_MAT	1983:01-2011:12	Δln
	Industrial Production Index - Manufacturing (Sic)	IP__MANUFACTURING	1983:01-2011:12	Δln
	Industrial Production Index - Residential Utilities	IP__RESIDENTIAL_UTILITIE	1983:01-2011:12	Δln
	Industrial Production Index - Fuels	IP__FUELS	1983:01-2011:12	Δln
	Napm Production Index	NAPM_PRODUCTION_INDEX	1983:01-2011:12	lv
	Capacity Utilization	CAPACITY_UTILIZATION	1983:01-2011:12	Δlv
GROUP II Labor Market	Index of Help-Wanted Advertising in Newspapers	INDEX_OF_HELP_WANTED_ADV	1983:01-2011:12	Δlv
	Ratio Help-Wanted Advertisings/ Unemployed	RATIO_OF_HELP_WANTED_ADS	1983:01-2011:12	Δlv
	Civilian Labor Force: Employed, Total	CIVILIAN_LABOR_FORCE__EM	1983:01-2011:12	Δln
	Civilian Labor Force: Employed, Nonagric. Industries	CIVILIAN_LABOR_FORCE_01	1983:01-2011:12	Δln
	Unemployment Rate: All Workers, 16 Years & Over	UNEMPLOYMENT_RATE__ALL_W	1983:01-2011:12	Δlv
	Unemployment By Duration: Average Duration in Weeks	UNEMP_BY_DURATION__AVERA	1983:01-2011:12	Δlv
	Unemployment By Duration: Persons Unemployed Less than 5 Wks	UNEMPLOY_BY_DURATION__PE	1983:01-2011:12	Δln
	Unemployment By Duration: Persons Unemployed 5 to 14 Wks	UNEMPLOY_BY_DURATION_01	1983:01-2011:12	Δln
	Unemploy. By Duration: Persons Unemployed more than 15 Weeks	UNEMPLOY_BY_DURATION_02	1983:01-2011:12	Δln
	Unemploy. By Duration: Persons Unemployed 15 to 26 Wks	UNEMPLOY_BY_DURATION_03	1983:01-2011:12	Δln
	Unemploy. By Duration: Persons Unemployed more than 27 Wks	UNEMPLOY_BY_DURATION_04	1983:01-2011:12	Δln
	Average Weekly Initial Claims, Unemploy. Insurance	INITIAL_CLAIMS_FOR_UNEMP	1983:01-2011:12	Δln
	Employees on Nonfarm Payrolls: Total Private	EMPLOYEES_ON_NONFARM_PAY	1983:01-2011:12	Δln
	Employees on Nonfarm Payrolls - Goods-Producing	EMPLOYEES_ON_NONFARM_01	1983:01-2011:12	Δln
	Employees on Nonfarm Payrolls – Mining	EMPLOYEES_ON_NONFARM_03	1983:01-2011:12	Δln
	Employees on Nonfarm Payrolls - Construction	EMPLOYEES_ON_NONFARM_04	1983:01-2011:12	Δln
	Employees on Nonfarm Payrolls - Manufacturing	EMPLOYEES_ON_NONFARM_02	1983:01-2011:12	Δln
	Employees on Nonfarm Payrolls - Durable Goods	EMPLOYEES_ON_NONFARM_06	1983:01-2011:12	Δln
	Employees on Nonfarm Payrolls - Nondurable Goods	EMPLOYEES_ON_NONFARM_05	1983:01-2011:12	Δln
	Employees on Nonfarm Payrolls - Service-Providing	EMPLOYEES_ON_NONFARM_07	1983:01-2011:12	Δln
	Employees on Nonfarm Payrolls - Trade, Transportation, and Utilities	EMPLOYEES_ON_NONFARM_08	1983:01-2011:12	Δln
	Employees on Nonfarm Payrolls - Wholesale Trade	EMPLOYEES_ON_NONFARM_09	1983:01-2011:12	Δln
	Employees on Nonfarm Payrolls - Retail Trade	EMPLOYEES_ON_NONFARM_10	1983:01-2011:12	Δln
	Employees on Nonfarm Payrolls - Financial Activities	EMPLOYEES_ON_NONFARM_12	1983:01-2011:12	Δln
	Employees on Nonfarm Payrolls - Government	EMPLOYEES_ON_NONFARM_11	1983:01-2011:12	Δln
	Employee Hours in Nonag. Establishments	INDEX_OF_AGGREGATE_WEEKL	1983:01-2011:12	Δln
	Avg Weekly Hrs of Prod or Nonsup Workers on Priv Nonfarm Payrolls - Goods-Producing	AVG_WEEKLY_HRS_OF_PROD_O	1983:01-2011:12	lv
	Avg Weekly Hrs of Prod or Nonsup Workers on Priv Nonfarm Payrolls - Mfg Overtime Hours	AVG_WEEKLY_HRS_OF_PRO01	1983:01-2011:12	Δlv
	Average Weekly Hours, Mfg.	AVERAGE_WEEKLY_HOURS__MF	1983:01-2011:12	lv
	Napm Employment Index	NAPM_EMPLOYMENT_INDEX	1983:01-2011:12	lv
GROUP III Housing	Housing Starts:Nonfarm(1947-58);Total Farm&Nonfarm(1959-)	HOUSING_STARTS_NONFARM_1	1983:01-2011:12	ln
	Housing Starts:Northeast	HOUSING_STARTS_NORTHEAST	1983:01-2011:12	ln
	Housing Starts:Midwest	HOUSING_STARTS_MIDWEST	1983:01-2011:12	ln
	Housing Starts:South	HOUSING_STARTS_SOUTH	1983:01-2011:12	ln
	Housing Starts:West	HOUSING_STARTS_WEST	1983:01-2011:12	ln
	Housing Authorized: Total New Priv Housing Units	HOUSING_AUTHORIZED__TOTA	1983:01-2011:12	ln
	Houses Authorized by Build. Permits:Northeast	HOUSES_AUTHORIZED_BY_BUI	1983:01-2011:12	ln
	Houses Authorized by Build. Permits:Midwest	HOUSES_AUTHORIZED_BY_01	1983:01-2011:12	ln
	Houses Authorized by Build. Permits:South	HOUSES_AUTHORIZED_BY_02	1983:01-2011:12	ln
	Houses Authorized by Build. Permits:West	HOUSES_AUTHORIZED_BY_03	1983:01-2011:12	ln

Appendix (continued)

List of (Macroeconomic) Variables Used to Extract Principal Components

Grouping	Variable	Acronym	Sample period	Notes
GROUP IV Consumption, Orders, Inventories	Purchasing Managers' Index	PURCHASING_MANAGERS__IND	1983:01-2011:12	lv
	Napm New Orders Index	NAPM_NEW_ORDERS_INDEX	1983:01-2011:12	lv
	Napm Vendor Deliveries Index	NAPM_VENDOR_DELIVERIES_I	1983:01-2011:12	lv
	Napm Inventories Index	NAPM_INVENTORIES_INDEX	1983:01-2011:12	lv
	Mfrs' New Orders, Consumer Goods and Materials	MFRS__NEW_ORDERS__CONSUM	1983:01-2011:12	Δln
	Mfrs' New Orders, Durable Goods Industries	MFRS__NEW_ORDERS__DURABL	1983:01-2011:12	Δln
	Mfrs' New Orders, Nondefense Capital Goods	MFRS__NEW_ORDERS__NONDEF	1983:01-2011:12	Δln
	Mfrs' Unfilled Orders, Durable Goods Indus.	MFRS__UNFILLED_ORDERS__D	1983:01-2011:12	Δln
	Manufacturing and Trade Inventories	MANUFACTURING_AND_TRADE_	1983:01-2011:12	Δln
	U. of Mich. Index of Consumer Expectations	U__OF_MICH__INDEX_OF_CON	1983:01-2011:12	Δlv
	Ratio, Mfg. and Trade Inventories to Sales	RATIO__MFG__AND_TRADE_IN	1983:01-2011:12	Δlv
GROUP V Money and Credit	Money Stock: M1	MONEY_STOCK__M1	1983:01-2011:12	Δ^2ln
	Money Stock:M2	MONEY_STOCK__M2	1983:01-2011:12	Δ^2ln
	Money Stock: M3	MONEY_STOCK__CURRENCY_HE	1983:01-2011:12	Δ^2ln
	Money Supply: M2 in 1996 Dollars	MONEY_SUPPLY__REAL_M2__A	1983:01-2011:12	Δln
	Monetary Base, Adj. for Reserve Requirement Changes	MONETARY_BASE__ADJ_FOR_R	1983:01-2011:12	Δ^2ln
	Depository Inst Reserves:Total, Adj. for Reserve Req Chgs	DEPOSITORY_INST_RESERVES	1983:01-2011:12	Δ^2ln
	Depository Inst Reserves:Nonborrowed,Adj. For Res Req Chgs	DEPOSITORY_INST_RESER01	1983:01-2011:12	Δ^2ln
	Commercial & Industrial Loans Oustanding in 1996 Dollars	COMMERCIAL_AND_INDUSTRIA	1983:01-2011:12	Δ^2ln
	Net Change Commercial & Indus Loans	CHANGE_IN_COMMERCIAL_AND	1983:01-2011:12	lv
	Consumer Credit Outstanding – Nonrevolving	CONSUMER_CREDIT_OUTSTAND	1983:01-2011:12	Δ^2ln
	Ratio, Consumer Installment Credit to Personal Income	RATIO__CONSUMER_INSTALLM	1983:01-2011:12	Δlv
GROUP VI Stock	S&P's Composite Common Stock: Dividend Yield	S_P_DIV_YIELD	1983:01-2011:12	Δlv
	S&P's Composite Common Stock: Price-Earnings Ratio	S_P_PE_RATIO	1983:01-2011:12	Δln
GROUP VII Exchange Rates	United States;Effective Exchange Rate	NOMINAL_E_ECTIVE_EXCHANG	1983:01-2011:12	Δln
	Foreign Exchange Rate: Switzerland (Swiss Franc per U.S.$)	FOREIGN_EXCHANGE_RATE__S	1983:01-2011:12	Δln
	Foreign Exchange Rate: Japan (Yen per U.S.$)	FOREIGN_EXCHANGE_RATE__J	1983:01-2011:12	Δln
	Foreign Exchange Rate: United Kingdom (Cents per Pound)	FOREIGN_EXCHANGE_RATE__U	1983:01-2011:12	Δln
	Foreign Exchange Rate: Canada (Canadian $ per U.S.$)	FOREIGN_EXCHANGE_RATE__C	1983:01-2011:12	Δln
GROUP VIII Prices	Producer Price Index: Finished Goods	PRODUCER_PRICE_INDEX__FI	1983:01-2011:12	Δ^2ln
	Producer Price Index: Finished Consumer Goods	PRODUCER_PRICE_INDEX_01	1983:01-2011:12	Δ^2ln
	Producer Price Index: Intermed Mat.Supplies & Components	PRODUCER_PRICE_INDEX_I_N	1983:01-2011:12	Δ^2ln
	Producer Price Index: Crude Materials	PRODUCER_PRICE_INDEX__CR	1983:01-2011:12	Δ^2ln
	Cpi-U: All Items	CPI_U__ALL_ITEMS	1983:01-2011:12	Δ^2ln
	Cpi-U: Apparel & Upkeep	CPI_U__APPAREL___UPKEEP	1983:01-2011:12	Δ^2ln
	Cpi-U: Transportation	CPI_U__TRANSPORTATION	1983:01-2011:12	Δ^2ln
	Cpi-U: Medical Care	CPI_U__MEDICAL_CARE	1983:01-2011:12	Δ^2ln
	Cpi-U: Durables	CPI_U__DURABLES	1983:01-2011:12	Δ^2ln
	Cpi-U: Services	CPI_U__SERVICES	1983:01-2011:12	Δ^2ln
	Cpi-U: All Items Less Food	CPI_U__ALL_ITEMS_LESS_FO	1983:01-2011:12	Δ^2ln
	Cpi-U: All Items Less Shelter	CPI_U__ALL_ITEMS_LESS_SH	1983:01-2011:12	Δ^2ln
	Cpi-U: All Items Less Medical Care	CPI_U__ALL_ITEMS_LESS_MI	1983:01-2011:12	Δ^2ln
	Pce, Implicit Price Deflator: Pce	PCE__IMPL_PR_DEFLATOR03	1983:01-2011:12	Δ^2ln
	Pce, Implicit Price Deflator:Pce, Durables	PCE__IMPL_PR_DEFLATOR01	1983:01-2011:12	Δ^2ln
	Pce, Implicit Price Deflator :Pce; Nondurables	PCE__IMPL_PR_DEFLATOR02	1983:01-2011:12	Δ^2ln
	Pce, Implicit Price Deflator:Pce; Services	PCE__IMPL_PR_DEFLATOR03	1983:01-2011:12	Δ^2ln
	Avg Hourly Earnings of Prod or Nonsup Workers on Priv Nonfarm Payrolls - Goods-Producing	AVG_HOURLY_EARNINGS_OF_P	1983:01-2011:12	Δ^2ln
	Avg Hourly Earnings of Prod or Nonsup Workers on Priv Nonfarm Payrolls – Construction	AVG_HOURLY_EARNINGS_O01	1983:01-2011:12	Δ^2ln
	Avg Hourly Earnings of Prod or Nonsup Workers on Priv Nonfarm Payrolls - Manufacturing	AVG_HOURLY_EARNINGS_O02	1983:01-2011:12	Δ^2ln
GROUP IX Others	Goldman Sachs Financial Conditions Index	GSERFCI_INDEX	1983:01-2011:12	
	Historical News-Based Policy Index	HISTORICAL_UNCERTAINTY	1983:01-2011:12	
	Liquidity Factor of Pastor and Stambaugh	LIQUIDITY_RISK_FACTOR	1983:01-2011:12	