ENDOGENOUSLY CLUSTERED FACTOR APPROACH TO MACROECONOMICS

OZGE SAVASCIN

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Economics.

Chapel Hill
2012

Approved by:

Neville R. Francis

Richard Froyen

Michael T. Owyang

Jonathan B. Hill

Eric Ghysels

ABSTRACT

OZGE SAVASCIN: Endogenously Clustered Factor Approach to Macroeconomics
(Under the direction of Neville R. Francis)

This dissertation constructs a novel factor approach to study the comovements of macroeconomic variables and introduces its two practical applications.

Factor models have become useful tools for studying international business cycles. Block factor models can be especially useful as the zero restrictions on the loadings of some factors may provide some economic interpretation of the factors. These models, however, require the econometrician to predefine the blocks, leading to potential misspecification. In my coauthored paper, we propose an alternative model in which the blocks are chosen endogenously. The model is estimated in a Bayesian framework using a hierarchical prior which allows series-level covariates to influence and explain how the series are grouped. Using similar international business cycle data as Kose, Otrok, and Whiteman (2005) we find our country clusters differ in important ways to those identified by geography alone. In particular, we find that similarities in institutions – e.g., legal systems, language diversity – may be just as important as physical proximity for analyzing business cycle comovements.

In another application, I use the endogenously clustered dynamic factor model to gain a better understanding of commodity price comovements and their determinants. From a large dataset of commodity prices I extract the fundamental sources behind the price dynamics and find that commodity price comovements are mostly the result of sparse cluster factors that represent correlations of distinct group of commodities. Endogenous clustering of these groups

does not represent the standard narrow classifications (indexes) of commodity prices as defined by statistical agencies (e.g. International Financial Statistics, Bureau of Labor Statistics). Characterization analysis on these factors identifies a wide range of macroeconomic variables like crude oil prices, fertilizer prices, and the federal funds rate as possible sources of commodity price comovements.

To My Mother, Father and Brother

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

CHAPTER 1

INTRODUCTION

In many aspects of science, natural grouping arises in many situations. An individual who lives in a certain location, in a certain country and belongs to a certain social group possesses characteristics distinct that come from that certain location, country and social group. Living styles for an individual may be similar to those who belong to the same social group even though they live in different locales. Likewise, different industries are affected by their nation's policies as well as by the things specific to their particular market.

The world consists of countries with similar legal backgrounds, language or geography. These aspects may present commonalities among economic measures such as gross domestic product, consumption, and prices. These economic measures with similar dynamics can form a cluster and may also have distinct properties compared to other groups of variables. In particular, similar dynamics (or comovements) of economic variables can be due to something general to the economy or due to something specific for some particular groups of variables. Specification of layers is the keystone in analyzing measures with these kinds of hierarchical structures. In my dissertation, I develop a model for specification of these layers or groups in investigating the comovements of macroeconomic variables.

One of the well known approaches to study comovements of macroeconomic variables that constitute a hierarchical structure is factor models. These models describe the variability of observed variables in terms of fewer measures called factors. The advantage of factor models is that the researcher could remain agnostic about the structure of the economy by treating these underlying sources, measures or factors, as unknowns. Therefore one does not need any tight assumptions or structural models to describe commonalities of variables. The drawback of basic factor models is that they neglect additional sources of fluctuations that may be off interest. Foerster et al. (2011) point this out in their paper where they investigate cross sectional industrial production data. They show that additional cross correlations of any kind could contaminate the global factor, and if not taken into account can lead to overestimation of the true nature and the contribution of common factors in explaining cross-product comovements.

A way to introduce these additional cross correlations is to form a block factor model as it provides a straightforward framework that allows a number of less pervasive factors (depending on the question these factors could be defined as regional or sector-specific) that account for possible correlations that are confined within particular groups of series. This technique however requires the knowledge of these groupings prior to the analysis at hand. For example, Kose Otrok and Whiteman (KOW, 2005) study international business cycles and define regions based on geographical configuration. Their model basically assumes that European countries form one group and their basic aggregates; gross domestic product, consumption and investment levels, move together due to a European factor. Likewise, all Latin American countries form another and share a common cycle; Mexico, USA and Canada are grouped as the North American cluster and are assumed to exhibit similar patterns, etc. On another paper, Kose Otrok and Prasad (2008) rely on income distributions to define the regions of a wider set of country series.

At a less aggregated level, variables such as prices may comove based on common factors or based on specific industry- or product-level factors. Industrial definitions are a natural way to put micro level price data into clusters. In stock price literature for example, it is highly common to place firms into clusters according to their SIC codes (King, 1966), firm sizes (Ng et al. (1992), Pindyck and Rotemberg (1993)), or simply according to industry affiliations (Brooks and Del Negro, 2005). Statistical agencies also report subgroups of price indexes based on industrial similarities. Consumer Price Index has food, beverages, energy and transportation listed as major subgroups. Commodity prices are grouped similarly as well. The reason we form indexes is to see what is going on within a market, and as argued by King (1966), a good index would be the one that is highly correlated within some set of products. However, some seemingly unrelated products can also be correlated (Pindyck Rotemberg, 1993). This may cause the market indexes to be contaminated by the spillovers from other industries; hence, they may no longer represent the general dynamics that are specific to that market. Given these kind of intercorrelations among variables, the factor models that use straightforward industry definitions to introduce industry specific factors may fall into potential misspecification which can alter their results significantly.

Overall, there are many ways a researcher can specify countries into "regions" or products into "industries". The natural question to ask is how we can be so sure that the way we group our data is the "best" in practice? Relying on the researcher to group the data is a tricky issue as different researchers will invariably have different criteria when forming these groups. One approach to avoid the researcher-biases with the group assignment would be to form every possible group and then employ a model selection criterion to determine the best cluster combination. However, with many time series this grid search procedure would be inefficient and possibly infeasible. The avenue in finding a model that gives us a tool to select clusters has not been taken within a factor framework. My dissertation attempts to fill this gap.

This dissertation introduces a novel approach to modeling comovement of cross-sectional economic variables and uses it to study International Business Cycles and commodity price dynamics. I model the key factors that cause distinct economic units to comove endogenously, in a purely data-driven way. Specifically, my approach to study synchronicity in economic variables refrains from imposing any prior belief on which groups of variables ought to comove and instead allow the data to form its own ex-post groupings. If indeed countries are grouped by continents than we should see those countries grouping together. Or if all food products share the same industry source, then data is free to form its clusters that way. The structure of the factor model is similar to the block factor models with the addition of a membership indicator determining which block a series belongs to. I estimate this indicator parameter using Bayesian techniques namely Metropolis Hastings within Gibbs sampling.

In my first paper, I study the synchronized dynamics of commodity prices. The past decade has witnessed a substantial comovement of commodity prices with different characteristics that raised several discussions and possible explanations in regards to what has been deriving these synchronized commodity price movements. A single price increase of a commodity would not be reflective on the overall economy but if the commodities exhibit persistent price increases all together, they can pass through to the core inflation rate and create a need for action by the monetary authorities.

In this paper, I aim to investigate how and why these commodity prices comove. The literature has largely focused on the second question and neglected the first one (Calvo (2008), Krugman (2008), Wolf (2008), Frankel (2008), Frankel and Rose (2009)). In order to know why, we should know how they comove first. Therefore, I take a step back and attempt to answer the first question using an endogenously clustered dynamic factor model that combines a range of factors (one global and several cluster factors) in a systematic way while identifying the group of commodities that are more likely to share cycles. After I extract the factors, I also try to understand what they are by additional Bayesian linear estimations to answer the question why commodity prices comove.

From a large dataset of non-energy commodity prices, I extract the fundamental sources behind the price dynamics and find that the commodity price comovements are mostly a result of sparse cluster factors that represent correlations of some groups of commodities. In particular, I find four main groups of products; "Timber", "Coffee", "Grains&Oils" and "Mixed". The latter of these is a highly correlated cluster of commodities consisting of seemingly unrelated products such as metals, agricultural materials and some food products. Additional analysis to characterize these correlations indicate the importance of federal funds rate, world demand, crude oil prices, and speculation in financial markets in deriving these common movements. Vegetable oils and grain prices react to the oil and fertilizer prices while timber industry seems to be significantly affected by the Chinese demand which could be related to the fact that Chinese buyers turned to the U.S. and Canada for wood after 2007 since Russia imposed higher tariffs on its logs.

In my coauthored paper, we study international business cycles by examining their relationships across countries. Characteristics such as industrial similarity, proximity, language, trade and coordinated monetary policies can lead some countries' business cycles to be correlated. Empirical models comparing business cycles generally take a block factor approach that builds on the assumption that countries within a block have cycles which are correlated through the block factor. As discussed previously, geographic proximity has been widely used to define county groupings. Perhaps sensible at a first glance, this is incomplete at best. For instance, U.S. and Mexican economies may not be better candidates for synchronous cycles just because they share a border.

In accord with the issues above, we develop a factor model and relax the assumption that blocks known ex ante. We utilize a Bayesian framework with hierarchical prior that helps us to incorporate possible macroeconomic measures which may influence how the series are grouped. We use measures to account for the legal and linguistic differences among countries, degree of openness, industrialization and trade. Using annual GDP data growth rates for 60 countries our findings suggest some evidence against the prevailing belief that geographic proximity is a major determinant of cross-country comovements. We find that one cluster represents a set of mostly industrialized nations (U.S., U.K., Australia, Canada, New Zealand, Denmark, and India). Outside of Denmark, one might interpret this cluster's comovement as a reflection of trade patterns perhaps spawned by commonalities attributable to being former British Commonwealth countries. Other regional clusters estimated in this paper suggest some geographic ties but there are still some countries that contradict this result. In particular, we identify a cluster which includes much of Europe and Japan; and another cluster which includes most of South America, Mexico, Norway, and Iceland. Our hierarchical prior covariate data suggests that linguistic diversity and legal institutions significantly determines these "regional" clusters.

The next sections present the complete papers.

CHAPTER 2

AN ENDOGENOUSLY CLUSTERED FACTOR APPROACH TO INTERNATIONAL BUSINESS CYCLES

The nature of business cycles is an issue central to macroeconomics. One way to better understand business cycles is to examine their relationships across countries, which has prompted several studies to consider common movements in business cycles across countries.[1] A related question asks what determines which countries share common movements in their business cycles. In particular, we might ask whether some characteristics (e.g., industrial similarity, proximity, language, trade) lead some countries' business cycles to be correlated.δ For example, Norrbin and Schlagenhauf (1996) estimated the role of industrial similarities in international business cycles but find a limited role for industry-specific shocks in explaining the forecast error variance of output across countries. Alternatively, coordinated (systematic) policies may be the impetus behind *any* synchronicity in business cycles across countries. McKinnon (1982) suggested coordinated monetary policies as a factor for synchronous cross-country business cycles.[2] Finally, correlation between macroeconomic aggregates across countries could be due to unobservable innovations – e.g., common international shocks or country-specific

[1] The relationship between business cycles across countries is not restricted to simple correlation. For example, Engle and Kozicki (1993) studied a number of common features across country pairs in the G7 and found common serial correlation. Clark (1998) and Clark and Shin (2000) find that region-specific shocks are important sources of comovement.

[2] This conclusion was reached after finding data consistent with the substitutability of national monies. In particular, McKinnon found that domestic money demand functions are unstable when no controls are made for foreign exchange rates. Additionally, in an empirical test of the role of borders in the synchronization of business cycles between US Census regions and across European countries Clark and Wincoop (2001) found limited roles for both monetary and fiscal policies.

shocks having spillover effects. Using structural vector autoregressions, Ahmed, Ickes, Wang, and Yoo (1993) conclude that spillovers from country-specific labor supply shock are more important than common shocks in generating international business cycles.

Empirical models comparing business cycles across countries generally take one of two approaches: (1) Country cycles are estimated separately and then compared or (2) Cycles are estimated jointly with numerous assumptions made on the correlation structure. For the most part, these approaches are motivated by the need to reduce complexity and potential parameter proliferation. The former leaves the country combinations unrestricted (i.e., any two countries' cycles can be correlated), while the latter explicitly excludes this. Which approach is taken can depend both on the question to be answered and the econometric techniques used to compute the cycle. For example, the first approach might define a country's cycle based on a Markov-switching or a trend-cycle decomposition, methods typically reserved for smaller systems of equations.[3] The second approach might define a common cycle via a factor model, where the factor loadings reflect the degree of correlation between country cycles [e.g., Bai (2003); Bai and Ng (2002); Forni, Hallin, Lippi, and Reichlin (2000, 2005); and Stock and Watson (2002a,b)].

In a series of recent papers, Kose, Otrok, and Whiteman (2003, 2008; henceforth KOW) propose a factor model with a block structure for the factor loadings.[4] This block structure provides a straightforward interpretation that may be lacking in standard factor models. Countries within a block have cycles which are correlated through a regional factor, while countries in different blocks are correlated only through a (set of) global factor(s). The standard factor model can emulate a block factor model if the loadings on the regional factors are close to

[3] Exceptions are Hamilton and Owyang (2009) and Kaufmann (2010) which use similar approaches to this paper in a Markov-switching environment.

[4] See also Boivin and Ng (2006); Onatski (2007); and Hallin and Liska (2008).

zero. Even in that case, however, the factors will produce some cross country correlation for countries outside its block. The significant advantage of the block factor model is that it allows a larger number of less pervasive (regional) factors, only a few of which affect any particular country. Thus, correlations across small numbers of countries may be identified in block factor models but missed in standard factor models in which the correlation is swamped by the large cross-section. The disadvantage of the block factor structure is that the blocks (or clusters) are predetermined, meaning we must make significant *ex ante* assumptions about which countries' cycles are correlated.

In this paper, we take the block factor approach but relax the assumption that the blocks are known *ex ante*. By being agnostic about block membership, we allow the data to cluster based on both their business cycle features and on country-specific characteristics. For example, countries could form groups based on their proximity, coordinated policies, and/or structural innovations. In this sense, we are not a priori guided by any one particular theoretical model. However, once the *ex post* country groupings are determined, potential commonalities within groups could aid in determining important features that any successful model of the international business cycle would need to possess. For example, if we find that common language is a better determinant of cross-correlation than physical distance, models of trade may consider language rather than geography as the determinant of iceberg costs.

The model has the block factor structure with an additional membership indicator determining which block a country belongs to. We assume block membership is a multinomial choice – i.e., a country cannot belong to more than one block. This multinomial approach to the block structure lends itself to estimation with Bayesian methods.[5] In the simplest execution of

[5] Our model has a similar flavor to the sparse factor model of Carvalho, Lopes, and Aguilar (2010).

the multinomial approach, we can assume either a uniform or Dirichlet prior on the membership indicator, giving the model the appearance of a clustering algorithm. For the uniform prior, cluster membership depends solely on the business cycle characteristics of the country's data as compared to the other members of the cluster. For the Dirichlet prior, the size of the cluster determines the *ex ante* probability a country is sorted to it. Another approach we explore is the use of a multinomial logistic prior on cluster membership [see also Frühwirth-Schnatter and Kaufmann (2008); Hamilton and Owyang (2009)]. The use of the logistic prior allows us to incorporate country-specific characteristics (e.g., location, industrialization, trade patterns) and enables us to test competing hypotheses about what influences which countries comove.

In Monte Carlo experiments with simulated data, we draw an obvious conclusion: empirical results, their economic interpretation, and the degree of confidence we place in them depend greatly on the specification of the block structure. For the case in which the clusters are known (and correct), the standard block factor model performs well. However, we find that small misspecifications of the block structure can lead to deviations from the true model and reductions in fit.

Our empirical application extends KOW's study of cross-country correlations. Using annual GDP data growth rates for 60 countries, we find that, although some regional/geographic correlation exists, there is also evidence against the prevailing belief that geographic proximity is the major determinant of cross-country comovements. We find evidence of only three clusters. The first consists of European countries excluding the U.K. and Denmark, along with Japan. A second cluster is composed of the U.K. and its former British Commonwealth countries: Australia, Canada, India, New Zealand, and the U.S., among others. A third cluster consists of South American countries, Mexico, and a few other countries. We find that – as opposed to physical distance – linguistic diversity and legal institutions are among the country-level determinants of this "regional" clustering. We also find that allowing the data to determine the clustering leads to higher contribution of the cluster (or regional) factor to the overall volatility of output.

The balance of the paper is as follows: Section presents the endogenous clustered factor model. Section outlines the Bayesian techniques we use to estimate the model. In this section, we focus on estimation of the model with a uniform prior on cluster membership. Section presents some Monte Carlo evidence showing how well our algorithm identifies the clusters and the consequences of exogenously misidentifying them. Section extends the model and the sampler with a multinomial logistic prior. Section presents results from the model with international business cycle data. Section summarizes and concludes.

1.1 Empirical Model

Suppose that we are presented with a panel of N series, $y_n = [y_{n1}, ..., y_{nT}]$, each of length T. We are interested in movements common across the series; these movements can be sorted into those which affect all series and those which affect only a few series. We will refer to the former as global factors and we will refer to the latter as cluster factors. Suppose there is but a single global factor and there are M clusters for which a series y_n belongs to a single cluster i.[6] That is, at each period, y_{nt} can be expressed as the sum of the global factor f_t; a single cluster factor F_{it}; an intercept β_{n0}; and an error term, ε_{nt}:

$$y_{nt} = \beta_{n0} + \beta_{nG} f_t + \beta_{ni} F_{it} + \varepsilon_{nt}, \tag{1}$$

[6] Increasing the number of global factors is straightforward. We discuss the choice of M below.

$i = 1, \ldots, M, \;\; t = 1, \ldots, T, \;\; n = 1, \ldots, N$ and $M \ll N$. Where β_{nG} and β_{ni} are the factor loadings.

The restriction that each series can belong only to one cluster is equivalent to zero restrictions on the factor loadings in a panel description of (1), giving it a block structure with which the factors can be "identified" as regions.[7] If we believe that some shocks are global – i.e., affect all of the series of interest – but some remain confined to the region or sector from which they originate, the model provides a framework with which we can perform regionally- or industrially-differentiated analysis [see Moench, Ng, and Potter (2009)]. In (1), we have *imposed* that series n belongs to cluster i, meaning that it is influenced by the ith cluster factor – in other words, a series' cluster is predetermined. But what if we are unsure which series should move together? KOW impose that the countries on the same continent comove; Moench, Ng, and Potter impose that sectoral data comove. While geographic proximity or industrial similarity may be a reason for two countries comovement, other causes – e.g., trade, demographics, level of industrialization – may also determine comovement. We, therefore, augment (1) to allow the clusters to be determined endogenously.

In endogenous clustering, the data choose the groupings. We define a cluster indicator, $\gamma_{ni} = \{0, 1\}$, which signifies whether series n belongs to cluster i, retaining the restriction that a series can belong only to a single cluster – i.e., $\sum_i^M \gamma_{ni} = 1$. Then, we have

$$y_{nt} = \beta_{n0} + \beta_{nG} f_t + \sum_i^M \gamma_{ni} \beta_{ni} F_t + \varepsilon_{nt}. \tag{2}$$

The model preserves the restrictions on the comovement of the series – series in different

[7] Exclusivity can be relaxed but would require modifications to the estimation algorithms presented below. These issues have been explored in other papers [e.g., Frühwirth-Schnatter and Lopes (2009)].

clusters comove only through the global factor, while series within the same cluster can comove apart from the global factor. However, in contrast to (1), we can now estimate the membership indicator, γ_{ni}, thereby allowing the data to determine the composition of the clusters.

We allow the error terms, ε_{nt}, to be serially correlated, following an AR(p_ε) process:

$$\varepsilon_{nt} = \psi_n(L)\varepsilon_{nt-1} + \epsilon_{nt},$$

where $\epsilon_{nt} \sim N\left(0, \sigma_n^2\right)$ and $E\left[\epsilon_{nt}\epsilon_{mt}\right] = 0$ for all $m \neq n$. The diagonality of the variance-covariance matrix implies that comovements between series not in the same cluster arise solely from the global factor.

Series within the same cluster, on the other hand, can comove via the global factor or the cluster factor. We assume that each factor (including the global factor) follows an AR(p_F) process of the form:

$$F_{it} = \phi_i\left(L\right)F_{it-1} + e_{it}, \tag{3}$$

where $\phi_i\left(L\right)$ is a polynomial in the lag operator and $e_{it} \sim N\left(0, \omega_i^2\right)$, where we normalize $\omega_i^2 = 1$ as is common in the literature.

1.2 Estimation

The endogenously clustered factor model outlined in the preceding section can be estimated using Bayesian techniques [see Gelfand and Smith (1990); Casella and George (1992); Carter and Kohn (1994)]. Bayesian methods allow us to estimate the cluster membership parameter directly using a single reversible jump Metropolis-Hastings step in the Gibbs sampler. In principle, one could estimate each cluster combination model using classical techniques and determine the final cluster composition via some model selection criteria. However, this would mean estimating and comparing a very large number of possible models. The optimal number of clusters (and, thus, the number of factors) are obtained by computing the marginal likelihoods for models with different numbers of clusters [see also Ghosh and Dunson (2008)].[8]

1.2.1 The Sampler

The sampler is an MCMC algorithm which draws from the conditional distributions of each parameter block conditional on the previous draws from the remaining parameters. The sequence of draws from the conditional distributions converges to the joint posterior. Let $\mathbf{Y}$ represent the data, $\boldsymbol{\Theta}$ represent the full set of model parameters, and $\mathbf{F}$ represent the full set of factors. The model parameters and factors can be drawn in five blocks: (1) the group membership indicators, $\boldsymbol{\gamma}$, jointly with the intercept and the factor loadings, $\boldsymbol{\beta}$, (2) the innovation variances, $\boldsymbol{\sigma}^2$; (3) the innovation autoregressive parameters, $\boldsymbol{\psi}$; (4) the factors, $\mathbf{F}$; (5) and the set of factor autoregressive parameters, $\boldsymbol{\phi}$. After initializing the sampler, the posterior distributions are computed with 10,000 iterations after 30,000 iterations discarded for convergence.

1.2.1.1 The Prior

For each series, the prior for factor loadings is normal, $\beta_n = [\beta_{n0}, \beta_{nG}, \beta_{ni}]' \sim N(\mathbf{b}_0, \mathbf{B}_0)$, and the innovation variances are inverse gamma, $\sigma_n^{-2} \sim \Gamma(\nu_0, \Upsilon_0)$. The factor and measurement error AR parameters also have normal priors, $\boldsymbol{\phi} \sim N\left(\mathbf{v}_0, \mathbf{V}_0^{-1}\right)$ and $\boldsymbol{\psi} \sim N\left(\mathbf{w}_0, \mathbf{W}_0^{-1}\right)$, respectively. As a first pass, we assume the cluster membership over all clusters is uniform –

[8] Marginal likelihoods are computed via the subsampling procedures proposed in Chib (1995) and Chib and Jeliazkov (2001). For more details, see the appendix.

that is, *a priori*, a series is equally likely to belong to any cluster. In section , we modify the sampler to incorporate country-specific characteristics into the cluster determination through a logistic hierarchical prior. The factors are assumed to have unit innovation variances. The hyperparameters for the prior distributions are given in Table 1. The draws of the variances and both sets of autoregressive parameters are straightforward and included in the Appendix.

1.2.1.2 Preliminaries

Before discussing conditional distributions for each block, it will be useful to specify a few key quantities. Let Ω_i represent the variance-covariance matrix of the stacked vector of p_F lags of the ith factor, which has elements given by

$$vec\left(\Omega_i\right) = \left(\mathbf{I} - \Phi_i \otimes \Phi_i\right)^{-1} vec\left(u_{p_F}' u_{p_F}\right),$$

where

$$\Phi_i = \begin{bmatrix} \boldsymbol{\phi}_i' & \\ \mathbf{I}_{p_F-1} & \mathbf{0}_{p_F-1\times 1} \end{bmatrix}$$

is the companion matrix associated with the ith factor, and u_{p_F} is a $(p_F \times 1)$ vector with a 1 as the first element and zeros as the rest. Define C_i as the Cholesky factor of Ω_i,

$$\Lambda_i = \begin{bmatrix} -\phi_{ip_F} & \cdots & -\phi_{i1} & 1 & 0 & \cdots & 0 \\ 0 & -\phi_{ip_F} & \cdots & -\phi_{i1} & 1 & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -\phi_{ip_F} & \cdots & -\phi_{i1} & 1 \end{bmatrix},$$

and

$$S_i^{-1} = \begin{bmatrix} C_i^{-1} & \mathbf{0} \\ & \Lambda_i \end{bmatrix}.$$

These quantities will be used to quasi-difference the factors. Similar quantities can be used to quasi-difference the data. For example, we could produce the analogue of S_i^{-1}, call it ζ_n^{-1}, for each series using the Cholesky factor of $\Sigma_n = (\mathbf{I} - \Psi_i \otimes \Psi_i)^{-1} vec\left(u_{p_\varepsilon}' u_{p_\varepsilon}\right)$ and the matrix Λ_n formed with the AR parameters for the error terms. Then, we can use S_i^{-1} and ζ_n^{-1} to quasi-difference the data and the factors.

1.2.1.3 Generating $\boldsymbol{\gamma}, \boldsymbol{\beta} | \boldsymbol{\Theta}_{-\boldsymbol{\gamma},\boldsymbol{\beta}}, \mathbf{F}, \mathbf{Y}$

For efficiency reasons, we draw β_n and γ_n jointly for each n. The joint draw of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ can be written as

$$q\left(\beta_n^*, \gamma_n^* | \boldsymbol{\Theta}, \mathbf{F}\right) = q\left(\gamma_n^* | \gamma_n, \boldsymbol{\Theta}, \mathbf{Y}, \mathbf{F}\right) \pi\left(\beta_n | \boldsymbol{\Theta}, \gamma_n^*, \mathbf{Y}, \mathbf{F}\right),$$

where we draw a candidate γ_n^* from $q\left(\gamma_n^* | \gamma_n, \boldsymbol{\Theta}, \mathbf{Y}, \mathbf{F}\right)$ which may or may not depend on the past (accepted) value of γ_n. Then, conditional on the candidate γ_n^*, we draw a candidate β_n^* from its full conditional distribution $\pi\left(\beta_n | \boldsymbol{\Theta}, \gamma_n^*, \mathbf{Y}, \mathbf{F}\right)$. This joint pair is then accepted or rejected.

Formally, let $\mathbf{X}_n^* = [\mathbf{1}_T, \mathbf{f}, \widetilde{\mathbf{F}}\gamma_n^*]$, where $\mathbf{1}_T$ is a $(T \times 1)$ vector of ones and $\widetilde{\mathbf{F}} = [\mathbf{F}_1, ..., \mathbf{F}_M]$ is the collection of cluster factors. Let $\overline{\mathbf{X}}_n^*$ and $\overline{\mathbf{Y}}_n^*$ represent the quasi-difference of $\mathbf{X}_n^*$ and $\mathbf{Y}_n$.[9] Then, the candidate β_n^* is drawn from

[9] See Chib and Greenberg (1994) for the details concerning the quasi-differencing procedure used here.

$$\beta_n | \mathbf{\Theta}_{-\beta_n,\boldsymbol{\gamma}_n}, \gamma_n^*, \mathbf{F}, \mathbf{Y} \sim N\left(\mathbf{b}_n^*, \mathbf{B}_n^*\right), \tag{4}$$

where $\mathbf{B}_n^* = \left(\mathbf{B}_0 + \sigma_n^{-2}\overline{\mathbf{X}}_n^{*\prime}\overline{\mathbf{X}}_n^*\right)^{-1}$, $\mathbf{b}_n^* = \mathbf{B}_n^*\left(\mathbf{B}_0^{-1}\mathbf{b}_0 + \sigma_n^{-2}\overline{\mathbf{X}}_n^{*\prime}\overline{\mathbf{Y}}_n^*\right)$.

Since we are drawing the β_n's from their full conditional densities – i.e., from $\pi\left(\beta_n^* | \gamma_n^*, \mathbf{\Theta}_{-\beta,\gamma}, \mathbf{F}, \mathbf{Y}\right)$, the value of β_n^* does not appear in the acceptance probability.[10] In this case, for each n, acceptance probability is

$$A_{n,\gamma} = \min\left\{1, \frac{\left|\mathbf{B}_n^*\right|^{1/2}}{\left|\mathbf{B}_n\right|^{1/2}} \frac{\exp\left(\frac{1}{2}\mathbf{b}_n^*\mathbf{B}_n^{*-1}\mathbf{b}_n^*\right)}{\exp\left(\frac{1}{2}\mathbf{b}_n\mathbf{B}_n^{-1}\mathbf{b}_n\right)} \frac{\pi\left(\gamma_n^*\right)}{\pi\left(\gamma_n\right)} \frac{q\left(\gamma_n | \gamma_n^*\right)}{q\left(\gamma_n^* | \gamma_n\right)}\right\}, \tag{5}$$

where $\mathbf{b}_n^*$ and $\mathbf{B}_n^*$ are defined as above and $\mathbf{b}_n$ and $\mathbf{B}_n$ are defined for γ_n, the value held over from the past draw.

To close this portion of the algorithm, we need to supply a proposal density for γ_n. We choose a symmetric density in which we draw a random element of γ_n and set this equal to one (setting all other elements equal to zero). The choice of the symmetric proposal makes the last term in (5) identically one.[11]

1.2.1.4 Generating $\mathbf{F}|\mathbf{\Theta}, \mathbf{Y}$

The set of factors are drawn recursively from the smoothed Kalman update densities using the techniques as described in Kim and Nelson (1999). However, the sign of the factors are not uniquely identified from the loadings – e.g., switching the signs on both a factor and its loading produces an observationally equivalent system. For identification, KOW normalize the

[10] For a formal proof of this assertion, see Appendix 1 in Troughton and Godsill (1997).

[11] Troughton and Godsill (1997) point out that the γ proposal density must allow some nonzero probability of revisiting the same model. That is, the probability that the candidate γ^* is equal to the last iteration's γ must be nonzero. If $\gamma^* = \gamma$, the acceptance probability is 1, but we still redraw β.

sign of the first factor loading in each group. Unlike KOW, we cannot restrict the sign of the first factor loading in each grouping as the clusters are not a priori known. We can, however, impose a sign on the first element (period 1) of each factor to resolve the sign identification issue. In some cases, this is not sufficient to avoid label switching (i.e., cases in which the sampler alternately draws F and $-F$). Thus, we also impose a normalization which selects either F or $-F$ depending on which is closest to the previous draw in mean squared distance. The draw of the factors is described in detail in the Appendix.

1.2.2 The Effect of Misspecification

Allowing the data to determine the clusters rather than setting them in advance highlights a tradeoff between the estimation uncertainty and potential misspecification. One would, therefore, want to evaluate the potential risks of each before proceeding with the difficult task of estimating the clusters. To this end, we perform a set of Monte Carlo (MC) experiments designed to determine how badly the clusters need be misspecified to outweigh the uncertainty of estimating them. Our MC experiments give the best chance to pre-specification by correctly setting the number of clusters – that is, the only source of potential misspecification is incorrectly assigning a series n to the wrong cluster.

We conduct 1000 MC replications by sampling 60 series of $T = 500$ evenly divided among 5 clusters. We begin by estimating the model with the (exogenous) correct cluster definitions and gradually increase the level of misspecification. We measure misspecification by the percentage of series exogenously allocated to the wrong cluster. Thus, 1.7% misspecification refers to one series allocated to the wrong cluster with all other series correctly specified. This format gives the exogenous model the best chance as incorrectly choosing the number of clusters would lead to obviously large amounts of misspecification. We then estimate the clusters endogenously and compute an entropy measure:

$$\mathbf{E} = \sum_{n=1}^{N} \left[\log(\boldsymbol{\sigma}_n^2) + \frac{\left(\overline{\mathbf{Y}}_n^* - \overline{\mathbf{X}}_n^* \beta_n\right)' \left(\overline{\mathbf{Y}}_n^* - \overline{\mathbf{X}}_n^* \beta_n\right)}{\boldsymbol{\sigma}_n^2} \right]$$

for each case. Higher entropy scores reflect poorer performance with relative entropy related to the familiar likelihood ratio statistic.[12]

Table 2 reports the results of the MC experiments. As expected, less misspecification is better than more misspecification. Interestingly, knowing the truth (zero misspecification) is statistically equivalent to estimating the truth (endogenous clustering), with the differences in the entropy scores likely due to variations in the small sample performances. Thus, we conclude that, in cases in which the truth is known, imposing the cluster composition is first best. However, if the cluster composition is not certain, allowing the data to determine the clusters reduces the risk of misspecification. It is important to keep in mind that, in these experiments, we know with certainty the true number of clusters. If the number of clusters is unknown, the potential for misspecification increases dramatically.

1.3 Incorporating Prior Beliefs of Cluster Membership

In the previous section, we assumed a flat prior over cluster membership. There are cases, however, for which prior information could be useful in characterizing the clusters. For example, similar industrial composition or geographic proximity could lead countries to respond to the same common factor. In this section, we consider an alternative logistic prior for the cluster membership indicator, γ_{ni}. For this multinomial prior, we include additional

[12] The entropy measure is calculated for each Gibbs iteration and the mean over all iterations is reported. Each MC replication is estimated with 40000 Gibbs iterations, with the first 30000 discarded for convergence.

blocks consisting of the hyperparameters δ and λ and the latent vector ξ. As in Hamilton and Owyang (2009), we can think of the prior hyperparameters as population parameters signifying the clusters' relationships.

1.3.1 Adding a prior for cluster membership

Suppose there exists a vector, $\mathbf{z}_{ni}$, of variables which may influence whether a series n belongs to cluster i. We assess the probability that series n belongs to cluster i as

$$\Pr\left[\gamma_{ni} = 1|\mathbf{z}_{ni}\right] = \begin{cases} \exp\left(\mathbf{z}_{ni}^{\prime}\boldsymbol{\delta}_i\right) / \left[1 + \sum \exp\left(\mathbf{z}_{ni}^{\prime}\boldsymbol{\delta}_i\right)\right] & i = 1, ..., M-1 \\ 1/\left[1 + \sum \exp\left(\mathbf{z}_{ni}^{\prime}\boldsymbol{\delta}_i\right)\right] & i = M \end{cases}, \tag{6}$$

for $n = 1, ..., N$ and where we have normalized $\boldsymbol{\delta}_M = 0$ for identification. In the multinomial framework, series n cannot be affiliated with more than one idiosyncratic cluster. Note also that the vector, $\mathbf{z}_{ni}$, need not be composed of the same variables for each cluster i. The standard approach to estimating the multinomial logistic is to augment the system in the spirit of Tanner and Wong (1987) with a latent vector that has the characteristic that the nonnegative element also reflects the cluster to which series n belongs. Formally, let $\boldsymbol{\xi}_i = (\xi_{1i}, ..., \xi_{Ni})'$ denote a set of latent vectors such that

$$\begin{array}{ll} \xi_{ni} \geq 0 & \text{if } \gamma_{ni} = 1 \\ \xi_{ni} < 0 & \text{otherwise} \end{array}. \tag{7}$$

Each ξ_{ni} can be thought of as a draw from a truncated logistic distribution. We follow Holmes and Held (2006) by defining a new latent variable, χ_{ni}, that will allow us to sample the hyperparameters of the priors along with the latent variables as additional Gibbs steps in the algorithm above.

Suppose that χ_{ni} has the limiting distribution of the Kolmogorov-Smirnov test statistic:

$$p\left(\chi_{ni}\right) = 8 \sum_{j=1}^{\infty} (-1)^{j+1} j^2 \chi_{ni} \exp\left(-2j^2 \chi_{ni}^2\right). \tag{8}$$

If $\chi_{ni} \sim KS$ and $o_{ni} \sim N(0,1)$, then $\xi_{ni} = \mathbf{z}_{ni}^{'}\boldsymbol{\delta}_i + 2\chi_{ni} o_{ni}$ has a logistic distribution with mean $\mathbf{z}_{ni}^{'}\boldsymbol{\delta}_i$ and unit scale parameter.[13] The cluster probabilities can be rewritten in terms of the new latent variables:

$$\Pr\left(\xi_{ni} > 0\right) = \frac{\exp\left(\mathbf{z}_{ni}^{'}\boldsymbol{\delta}_i\right)}{1 + \sum_{j=1}^{M-1} \exp\left(\mathbf{z}_{nj}^{'}\boldsymbol{\delta}_j\right)}.$$

The following subsections demonstrate how to draw the hyperparameters governing the cluster prior probabilities.

1.3.2 Augmenting the Sampler

The sampler outlined in Section can be augmented to account for the logistic prior described above. Conditional on the γ_{ni}'s, draws of most of the model parameters remain unchanged. The change to the logistic prior does alter the acceptance probability in the joint draw of γ_{ni} and β_{ni} to the probability defined by (6). The only other modification is in the form of two additional blocks sampling the three prior parameters: covariate effects, δ; the logistic variances, λ; and the vector of latent variables, ξ. Each of these blocks is drawn by iterating (jointly) over the $M-1$ unnormalized clusters.

[13] See Devroye (1986).

1.3.2.1 Generating $\delta|\boldsymbol{\Theta}, \boldsymbol{\xi}, \boldsymbol{\lambda}, \mathbf{F}, \mathbf{Y}$

Conditional on $\boldsymbol{\xi}$ and $\boldsymbol{\lambda}$, $\boldsymbol{\delta}_i$ are the slope coefficients from a standard Normal regression model for each of the form:

$$\boldsymbol{\xi}_i = \mathbf{Z}_i \boldsymbol{\delta}_i + \mathbf{v}_i,$$

where $\mathbf{Z}_i = [\mathbf{z}_{1i}, ..., \mathbf{z}_{Ni}]'$, $\mathbf{v}_i \sim N(\mathbf{0}, \boldsymbol{\tau}_i)$, and $\boldsymbol{\tau}_i = diag[\lambda_{1i}, ..., \lambda_{Ni}]$. We assume a normal prior for the logistic slope parameters, $\boldsymbol{\delta}_i \sim N(\mathbf{d}_i, \mathbf{D}_i)$. Thus, the covariate effects can be drawn from the posterior $\boldsymbol{\delta}_i|\mathbf{Y}, \boldsymbol{\Theta}, \mathbf{F} \sim N(\mathbf{d}_i^*, \mathbf{D}_i^*)$, where

$$\mathbf{d}_i^* = \left(\mathbf{D}_i^{-1} + \mathbf{Z}_i' \boldsymbol{\tau}_i^{-1} \mathbf{Z}_i\right)^{-1} \left(\mathbf{D}_i^{-1} \mathbf{d}_i + \mathbf{Z}_i' \boldsymbol{\tau}_i^{-1} \boldsymbol{\xi}_i\right)$$

and

$$\mathbf{D}_i^* = \left(\mathbf{D}_i^{-1} + \mathbf{Z}_i' \boldsymbol{\tau}_i^{-1} \mathbf{Z}_i\right)^{-1}.$$

1.3.2.2 Generating $\boldsymbol{\xi}$ and $\boldsymbol{\lambda}|\boldsymbol{\Theta}, \boldsymbol{\delta}, \mathbf{F}, \mathbf{Y}$

If we condition on λ_{ni}, then ξ_{ni} would be Normal, $\xi_{ni}|\boldsymbol{\delta}_i, \lambda_{ni} \sim N(m_{ni}, \lambda_{ni})$, for the $i = 1, ..., M-1$ unnormalized clusters. The mean of the Normal distribution reflects this normalization:

$$m_{ni} = \mathbf{z}_{ni}' \boldsymbol{\delta}_i$$

Without that conditioning but given γ_{ni}, ξ_{ni} is a truncated logistic with mean m_{ni}. The truncation point is at zero, where γ_{ni} determines the direction of the truncation: $\xi_{ni} \geq 0$ if $\gamma_{ni} = 1$ and $\xi_{ni} < 0$ if $\gamma_{ni} = 0$.

Then, to sample λ_{ni}, Holmes and Held (2006) suggest that we can draw a candidate $\widehat{\lambda}_{ni}$

from a Generalized Inverse Gaussian distribution. The candidate, $\hat{\lambda}_{ni}$, is accepted or redrawn based on the algorithm described by Holmes and Held (2006).

1.4 Empirical Application

As an empirical application, we reconsider the model proposed in KOW in which geography is the sole determinant of cross-country comovements. We include in the hierarchical prior sets of variables which have been suggested to affect trade between countries. In doing this, we can assess the sources of business cycle comovements.

1.4.1 Data

Our measure of business cycle activity is the annual constant-price chain-weighted real GDP growth rate (computed as the difference in the log of real GDP) taken from the 6.3 version of the Penn World Tables [Heston, Summers, and Aten (2009)].[14] To maintain comparison, we choose the same 60 countries located in seven regional blocks from KOW.[15]

In addition to the real GDP data, the use of the logistic prior requires covariate data, $\mathbf{Z}_i$. Our covariate dataset includes domestic and international variables as well as indices of institutional differences. We will focus on the differences in legal and linguistic institutions. We have a total of seven covariates that inform the logistic prior: (1) The degree of economic openness, defined as the ratio of imports and exports to GDP; (2) Investment share of real GDP; (3) An index of conflict resolution and sophistication of the legal system as captured

[14] KOW's business cycle data include other series in addition to real GDP, allowing them to estimate country factors. We focus on the comovements across countries by restricting the model to a single business cycle indicator. Extension to include country factors is left for future research.

[15] To increase number of annual observations, we use a later version of the PWT. Ponomareva and Katayama (2010) discuss the hazards of comparing empirical studies across versions of the PWT. Table 15 in the data appendix shows the 60 countries in the estimation along with the regional groupings imposed in KOW.

by the manner in which lower courts facilitate landlords' collection of checks (and remedies for bounced checks); (4) An index of language diversity within each country; (5) An index of production dispersion relative to the rest of the world; (6) An index of export dispersion from each country's exporting partners; and (7) A similar index of import dispersion from each country's importing partners. The covariate data are summarized in Table 3.

Openness measures the size of trade as a fraction of GDP. This variable proxies the extent of a country's dependence on foreign economies and exposure to external shocks, without controls for the types of goods traded or the identities of trading partners, allowing us to determine whether countries cluster based on the (relative) extent of their (direct) exposures to international shocks. Investment share of GDP is meant to capture the degree of industrialization; similar levels of industrialization may make countries susceptible to similar shocks inducing comovements.

The indices in (3) and (4) are included to test the extent to which institutions matter for clustering. Our institutional variables are the level of formality of the civil-court system and the degree of linguistic-diversity. Djankov et al. (2003) construct the lower court system's *formalism* index in (3) which *"measures substantive and procedural statutory intervention in judicial cases at lower-level civil trial courts [p.469]"*. We hypothesize that trade flow between countries with similar conflict resolution processes in civil courts could be higher as individuals may prefer to form relationships in countries with familiar legal set-ups.

The ethnolinguistic index in (4) is taken from La Porta et al. (1999) and measures the degree of language diversity, the probability that two randomly selected individuals in a given country speak different languages, are not speaking the official language, or are not speaking the most widely used language.

Finally, Baxter and Kouparitsas (2003) construct the indices in (5) - (7) to analyze how the composition of a country's production and trade differ from the rest of the world and its trading partners. These indices are akin to variance measures, with the exception that the export and import dispersions are weighted by sectoral export and import shares. A look at the trade dispersion indices, (6) and (7), reveals that they capture both the strengths of trading relations with different countries and the strength in the diversity of goods traded.[16] Baxter and Kouparitsas find that industrialized nations have dispersions similar to the rest of the world (the average country) for all three indices while developing countries systematically have higher values of dispersions. On the trade side, this is consistent with the fact that the bulk of trade of an industrialized nation is with other industrialized nations, while developing nations have trade relations more evenly spread across developed and developing nations. By including these indices, we are allowing for the possibility that countries cluster on the similarities in their production structures (in terms of types of goods produced) and/or on the compositions of their trade (both in terms of types of goods traded and the trading partners).

1.4.2 Results

We first determine the optimal number of country clusters which, for simplicity, we compute with a flat hierarchical prior on cluster membership. This allows us to determine the optimal number of clusters based solely on the business cycle properties of GDP. With flat model priors, the Bayes factors are identically the posterior odds. Table 4 presents these results. The model with the highest probability is the model with three clusters. Two and six clusters have the next highest marginal likelihood; however either alternative require more than 100 times higher prior likelihood to be preferred. The model with seven clusters – the specification which nests the one estimated by KOW – has one of the lowest likelihood of the

[16] We refer the reader to the data appendix for more details about the construction of these indices.

alternatives tested.[17]

We now estimate the model using the logistic prior for the specification with three regional factors and one global factor. Figure 1 plots the median of the global factor along with its 16th and 84th percentiles; the shaded areas show the NBER-defined recession dates defined as a year in which any quarter was in recession. While the NBER recessions are defined only for the U.S., they serve as reference points. The global factor roughly represents a global cycle with most countries' factor loadings being negative. The global factor spikes around 1975, 1982, 1998, and 2001. With the exception of 1998, these periods are roughly associated with U.S. NBER recessions.

Figure 2 shows the first cluster factor along with its 68 percent probability bands and the NBER recessions. Figure 3 shows the posterior inclusion probabilities for this cluster. Dark blue indicates countries which are very likely to be included in this cluster. Yellow indicates countries which are very likely not associated with the cluster. Countries in white are not included in our sample. Note, in particular, that this cluster does appear to demonstrate some regional/geographic properties. The cluster includes, with high probability, Japan and many of the countries in Europe. Some European countries – e.g., Iceland and Ireland – belong with more than 50 percent probability. Brazil, Thailand, and Pakistan also belong with more than 50 percent probability. Not all of the European countries, however, appear to belong to this cluster. In particular, the U.K. and Denmark are excluded.

Figure 4 shows the second cluster factor. This factor clearly appears to decline around NBER recessions. Figure 5 shows why. The U.S. belongs to this cluster with probability 1; the cluster also includes Australia, Canada, Hong Kong, India, Malaysia New Zealand, and the

[17] In this case, the algorithm chooses nearly empty clusters at some Gibbs iterations, suggesting that seven clusters far exceeds the optimal number.

U.K. with very high posterior probability. Also included in this cluster are Denmark and many of the sub-Saharan African countries including South Africa.

Figure 6 shows the final factor and Figure 7 shows the composition of its cluster. Again, the cluster displays some regional/geographic characteristics with some notable exceptions. The cluster includes with high probabilities most of the countries in South America, with the exception of Brazil. Mexico, the Philippines, and a few African countries also belong with high probability.

As opposed to a purely continental approach such as KOW, our results suggest that a country like Mexico is much more likely to belong have similar cycles to its common language South American neighbors than its more geographically proximate neighbor, the U.S. These results suggest that common culture – either through linguistic or legal similarities – matter more for cyclical commonality than iceberg costs usually associated with geographic proximity. Table 5 shows the posterior means for the logistic covariates along with the 16th and 84th percentiles of the posterior distributions. The level of industrialization proxied by the country's investment share of GDP is important in determining the clusters. Also, similarities in the countries' legal systems and in their linguistic diversity also appear relevant. This view is consistent with the notion that trade flows – and, therefore, business cycle comovements – are more likely across countries with similar institutions.

One measure that can jointly capture the importance of both the factor and its loading can be obtained through a variance decomposition. Table 6 shows the percentage of each country's output volatility attributable to the global and regional factors and the idiosyncratic shock. While the results are, again, not directly comparable to KOW, there are a number of qualitative similarities and differences that highlight the effect of estimating the clusters. KOW find that, in general, the global factor explains a greater portion of the volatility in the more industrialized countries. Moreover, they conclude that the regional factors explain only a very small portion of macroeconomic fluctuations (about 3.6 percent on average of the 60 countries' output fluctuations). Our results suggest that there exists a much larger role for the "regional" factor if region is estimated by the countries' cyclical commonality. In fact, our cluster factors explain an average of 22.5 percent of the countries' output fluctuations.

There are a few reasons this difference may not be surprising. First, KOW's regional factors are defined as the common component for three series for each country. The inclusion of the additional two macroeconomic series could potentially contaminate their regional factor's ability to explain output fluctuations. Second, imposing rather than estimating the regions may lead to the same misspecification discussed in the Monte Carlo experiments above. When countries are included in a region with countries which it does not actually share a common factor, the factor and the associated loadings may be biased.

Indeed, when the model is estimated with only output with KOW cluster definitions, the difference between the average variances explained by the regional factors in the two models is not as large, about 1.2 percentage points. The variance explained by the global factor in the exogenous model is about 4 percentage points lower. The largest difference, however, comes from the countries in the former British Commonwealth. In the purely geographic model which would place these countries in three separate regions, the regional factor would explain 36 percent of the variation in output for these countries (Australia, Canada, New Zealand, the U.K., and the U.S.). In the endogenous model which groups them together, the regional factor explains 57 percent of their output variation. This increase in explanatory is important, especially given that these countries account for a substantial share of the total output of the 60 countries in the sample.

1.5 Conclusions

A great deal of research has been done on measuring the comovement of business cycle variables across countries. Limited by the potential proliferation of the estimated parameters, these empirical models typically (1) compare business cycles which are estimated country-by-country; (2) use models of relatively few countries (e.g., bilateral analyses); and/or (3) impose the structure of the correlations *ex ante*. One application of the latter, KOW, estimates a factor model in which the correlation structure across countries is assumed to be determined by geographic proximity – that is, countries which share a continent also share a common unobserved factor.

In this paper, we allow the data to determine which countries share common factors. Our model allows for a number of possible alternative country characteristics which can affect how countries are grouped. In Monte Carlo experiments, we show that misspecifying the regions can have consequences on the fit of the model. In the data, we find evidence that common geographic region is a component but not the only determinant of the country groupings. These results, then, verify some of the underlying reasoning behind KOW's selection of shared continent as the basis of defining a region. However, while there do appear to be some localized comovements (e.g., South America, and Europe), these comovements stretch beyond what would be narrowly considered geographic regions and exclude some countries which would ordinarily be associated by continent. In particular, continental Europe appears to share a common cyclical component with Japan but not with the U.K. and the bulk of South America appears to share a cycle with Mexico but less so with Brazil. One cluster consisting of the U.S., U.K., and some other former British Commonwealth countries belies geography or proximity as the driving force behind the cyclical commonality and suggests other fundamental forces linking the countries.

1.6 Tables and Figures

Priors for Estimation - I

Parameter	Prior Distribution	Hyperparameters	
β_n	$N(\mathbf{b}_0, \mathbf{B}_0)$	$\mathbf{b}_0 = \mathbf{0} \times \mathbf{I}_3$; $\mathbf{B}_0 = \mathbf{I}_3$	$\forall n$
σ_n^{-2}	$\Gamma\left(\frac{\nu_0}{2}, \frac{\Upsilon_0}{2}\right)$	$\nu_0 = 6$; $\Upsilon_0 = 0.1$	$\forall n$
γ_n	$\mathbf{U}(\kappa_0)$ or *Logistic*	$\kappa_0 = \frac{1}{M}$	$\forall n$
$\boldsymbol{\phi}$	$N(\mathbf{v}_0, \mathbf{V}_0)$	$\mathbf{v}_0 = \mathbf{0}_{p_F}, \mathbf{V}_0 = \frac{1}{2}\mathbf{I}_{p_F}$	$\forall i$
$\boldsymbol{\psi}$	$N(\mathbf{w}_0, \mathbf{W}_0)$	$\mathbf{w}_0 = \mathbf{0}_{p_\varepsilon}, \mathbf{W}_0 = \frac{\mathbf{1}}{\mathbf{2}}\mathbf{I}_{p_\varepsilon}$	$\forall n$
δ_i	$N(\mathbf{d}_0, \mathbf{D}_0)$	$\mathbf{d}_0 = \mathbf{0} \times \mathbf{I}_7$; $\mathbf{D}_0 = \mathbf{2} \times \mathbf{I}_7$	

Table 1: Priors for Estimation - I. Notes: n denotes the series. i indicates the cluster, where M is the total number of cluster factors. p's are the maximum number of lags in the error and factor lag polynomials.

Cluster Misspecification

	60%	40%	20%	6.7%	5%	3.4%	1.7%	*None*	*Endo*
Entropy	3372.2	3339.4	3302.7	3299.76	3295.30	3291.49	3289.98	3288.80	3287.45

Table 2: Cluster Misspecification. Notes: The table reports the median for 1000 Monte Carlo replications with sample size of 50 periods. Unless otherwise specified, each sample contains 60 series, 5 cluster factors, and 1 global factor. The panel compares the results from the endogenous cluster algorithm to the exogenous cluster algorithm for different degrees of missepcification. The column headers indicate the percent of the series in the exogenous clusters are misallocated. 'None' specifies the exogenously clustered model with known clusters. 1 misallocated series out of 60 equates to 1.66 percent misspecification, etc.

Covariate Data

Purpose	Variable	Mnemonic
Trade	Openness	OPEN
Industrialization	Investment Share of GDP	KI
Formalism Index	Collection of Bounced Checks	CHECK
Linguistic Diversity within a Country	Ethnolinguistic Fraction	LIN
Production Dispersion	Production Dispersion versus World	ProDisp
	Export Dispersion versus Export Partners	ExDisp
	Import Dispersion versus Import Partners	ImDisp

Table 3: Covariate Data.

Model Choice

	$\ln f(\mathbf{Y}\vert\mathbf{\Theta}^*)$	$\ln \pi(\mathbf{\Theta}^*)$	$\ln \widehat{\pi}(\mathbf{\Theta}^*\vert\mathbf{Y})$	$\ln \widehat{m}(\mathbf{Y})$	Odds
$f=2$	-6772	-2173	211	-9157	-58
$f=3$	**-6746**	-2193	159	**-9099**	0
$f=4$	-6893	-2206	114	-9214	-115
$f=5$	-6999	-2221	91	-9312	-213
$f=6$	-6879	-2229	12	-9121	-22
$f=7$	-7008	-2236	55	-9301	-202

Table 4: Model Choice. Notes: The table shows the log marginal likelihood for model with various numbers of clusters estimated with the empirical data. The third column shows the difference in the log marginal likelihoods between the best model and each other model. The last column shows how much more likely the best model is compared to each other model.

Logistic Coefficients

Variable	Cluster 2	Cluster 3
OPEN	-0.19	0.46
	(-0.77 0.39)	(-0.11 1.01)
KI	-0.49	**-1.21**
	(-1.12 0.14)	(-1.82 -0.54)
CHECK	**-1.86**	**2.47**
	(-2.47 -1.24)	(1.82 3.09)
LIN	**0.99**	**-0.89**
	(0.29 1.71)	(-1.48 -0.31)
ProDisp	**-1.61**	**-1.24**
	(-2.89 -0.34)	(-2.48 -0.03)
ExDisp	**-1.47**	-0.93
	(-2.82 -0.16)	(-2.26 0.41)
ImDisp	**-1.90**	**-1.42**
	(-3.21 -0.61)	(-2.76 -0.07)

Table 5: Logistic Coefficients. Notes: Posterior means are reported for each of the covariate variable in clusters 2 and 3. The first cluster (cluster 1) covariate coeffients are normalized to zero. Values in bold indicate coefficients for which zero is not within the 68-percent coverage interval. The numbers in parentheses are the 16th and 84th percentiles of the posterior distributions.

Variance Decompositions - Posterior Means

Country	Global	Cluster	Idio.	Country	Global	Cluster	Idio.
Argentina	0.1	21.7	78.2	Japan	22	8.9	69.1
Australia	6.5	49.7	43.8	Kenya	0	4.4	95.6
Austria	21.4	49.2	29.3	Korea	60.3	0.9	38.7
Bangladesh	0.6	0.1	99.3	Luxemburg	7.9	36.9	55.2
Belgium	21.6	66.2	12.2	Malaysia	72.3	0.7	27
Bolivia	3.1	19.5	77.4	Mexico	7.1	19	73.8
Brazil	5	24.8	70.3	Morocco	0	0.2	99.8
Cameroon	7.4	3.7	88.9	Netherlands	12.4	44.3	43.3
Canada	6.5	77.9	15.6	New Zealand	0.9	25.2	73.9
Chile	20.6	11.8	67.6	Norway	3.9	21.8	74.2
Colombia	22.2	27.3	50.5	Pakistan	0.2	11.6	88.2
Costa Rica	0.9	28	71.2	Panama	3.9	9.3	86.8
Ivory Coast	0.7	2.5	96.8	Paraguay	0	7.8	92.2
Denmark	9.6	33.4	57	Peru	0.6	16.5	82.9
Dom. Republic	4.7	10.2	85.1	Philippines	11.9	15	73.1
Ecuador	1.1	25.5	73.5	Portugal	23.8	34.3	42
El Salvador	0.6	23.1	76.3	Senegal	15.7	0.1	84.2
Finland	5.5	40.8	53.7	Singapore	55.1	0.3	44.5
France	12.5	73.2	14.3	South Africa	1	0.1	98.9
Germany	31.9	16.7	51.4	Spain	0.7	22.1	77.2
Greece	14	18.5	67.5	Sri Lanka	5.2	16.9	77.9
Guatemala	30.9	18.1	51	Sweden	8.6	59.6	31.8
Honduras	0.5	19.3	80.2	Switzerland	4.8	69.9	25.3
Hong Kong	36.4	7	56.5	Thailand	57	3.9	39.2
Iceland	0	14.1	85.9	Trinidad & Tobago	10.2	5.7	84.1
India	1.5	7.1	91.4	United Kingdom	14.5	52.5	32.9
Indonesia	58.8	0.3	40.9	United States	24.6	64.3	11.1
Ireland	21.8	14.6	63.6	Uruguay	2.3	18.6	79.1
Italy	22.6	53.4	24	Venezuela	0.4	21	78.5
Jamaica	2.5	0	97.5	Zimbabwe	0	0.6	99.4

Table 6: Variance Decompositions - Posterior Means. Notes: The table summarizes the variance decompositions in percentages. Each row shows the variation in GDP growth that is attributable to the global, cluster and idiosyncratic factors. In calculation of the variance share of clusters, members are assumed to belong to a cluster if they pick the said cluster majority of the Gibbs run. In other words, the modal values for the indicator function is used to determine which cluster a country is in, then the variance attributable to that specified cluster is calculated.

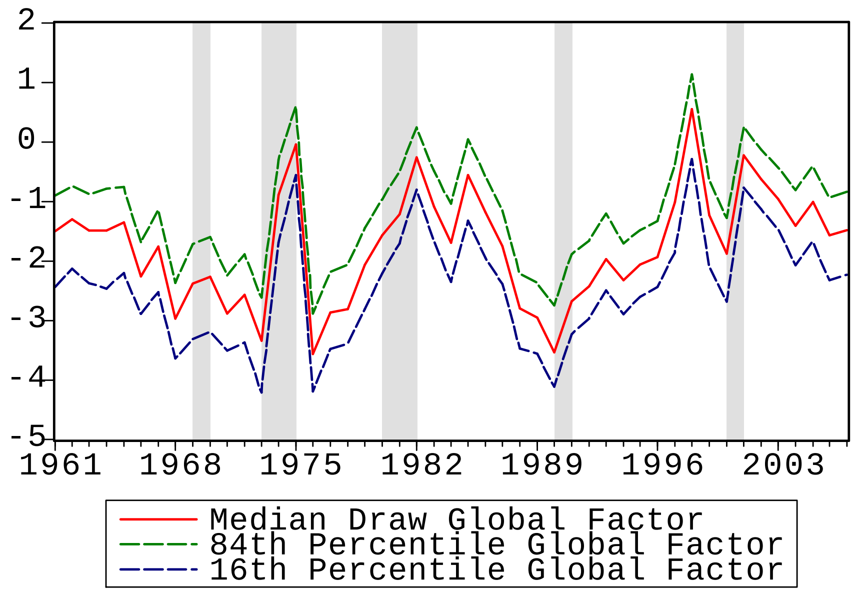

Figure 1: Global Factor. Notes: The red solid line is the median of the posterior distribution of the global factor. Dashed lines represent the 16th and 84th percentiles. Shaded regions are annual NBER recessions, where a recession is defined as a year in which any quarter was in recession according to the Business Cycle Dating Committee turning points.

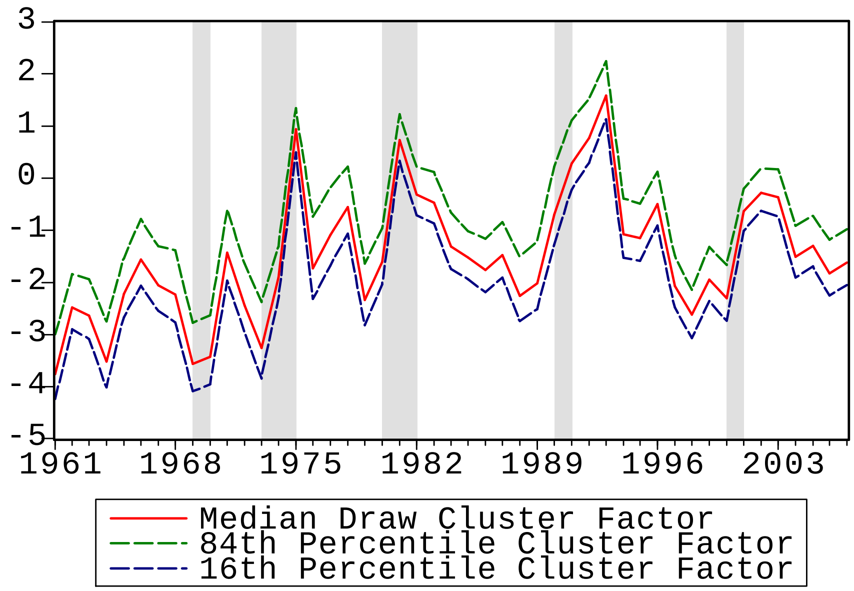

Figure 2: Cluster 1 Factor. Notes: The red solid line is the median of the posterior distribution of Cluster 1's factor. Dashed lines represent the 16th and 84th percentiles. Shaded regions are annual NBER recessions, where a recession is defined as a year in which any quarter was in recession according to the Business Cycle Dating Committee turning points.

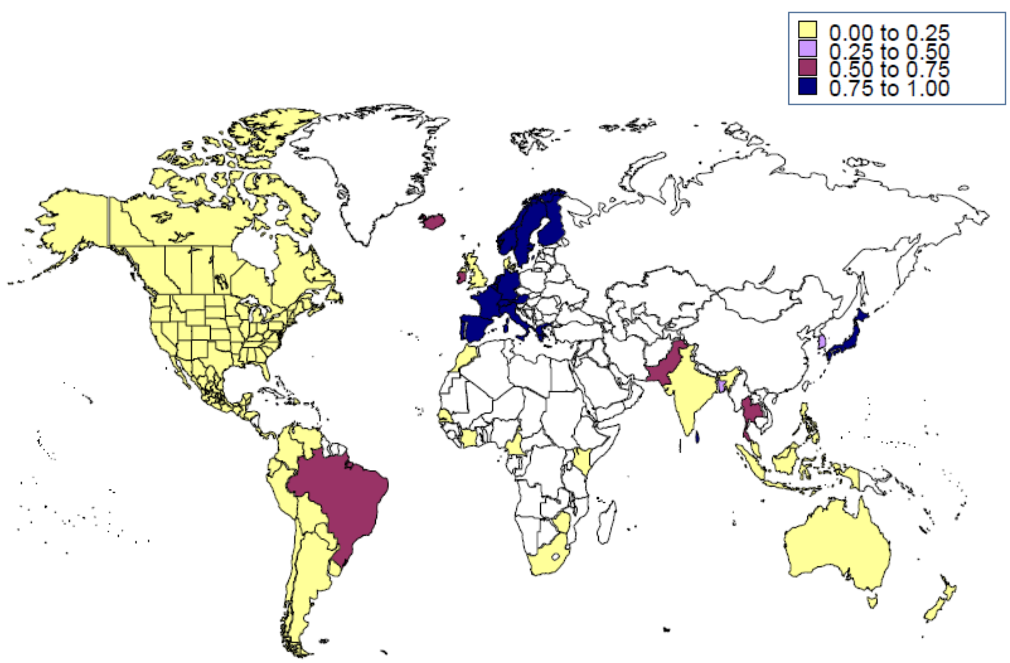

Figure 3: Cluster 1 Composition. Notes: The map shows the posterior probabilities of countries included in Cluster 1. Countries in white are omitted from the sample.

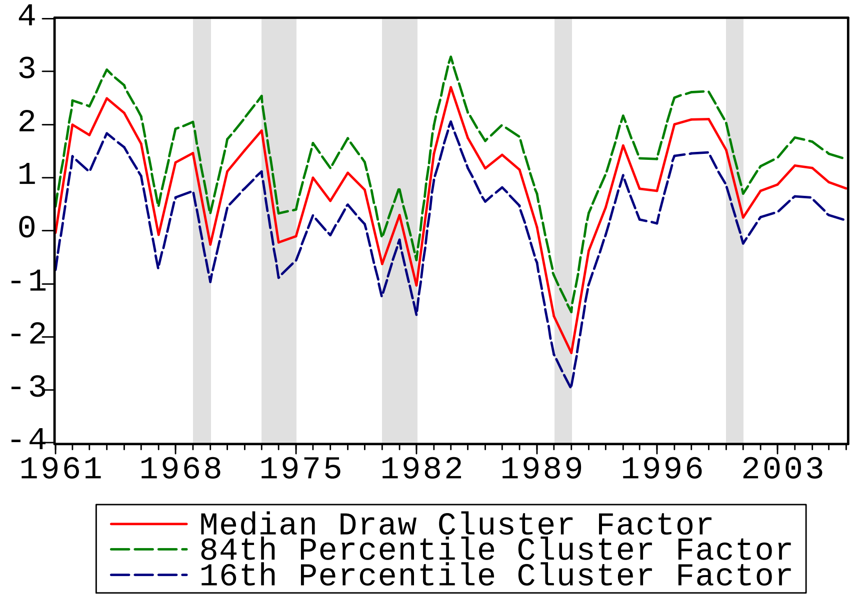

Figure 4: Cluster 2 Factor. Notes: The red solid line is the median of the posterior distribution of Cluster 2's factor. Dashed lines represent the 16th and 84th percentiles. Shaded regions are annual NBER recessions, where a recession is defined as a year in which any quarter was in recession according to the Business Cycle Dating.

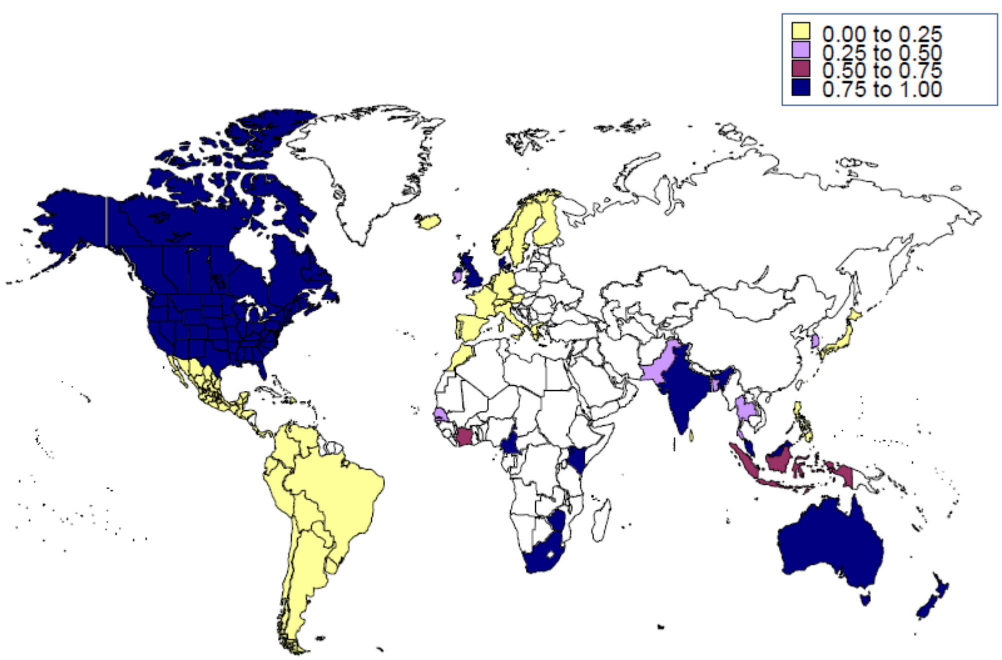

Figure 5: Cluster 2 Composition. Notes: The map shows the posterior probabilities of countries included in Cluster 2. Countries in white are omitted from the sample.

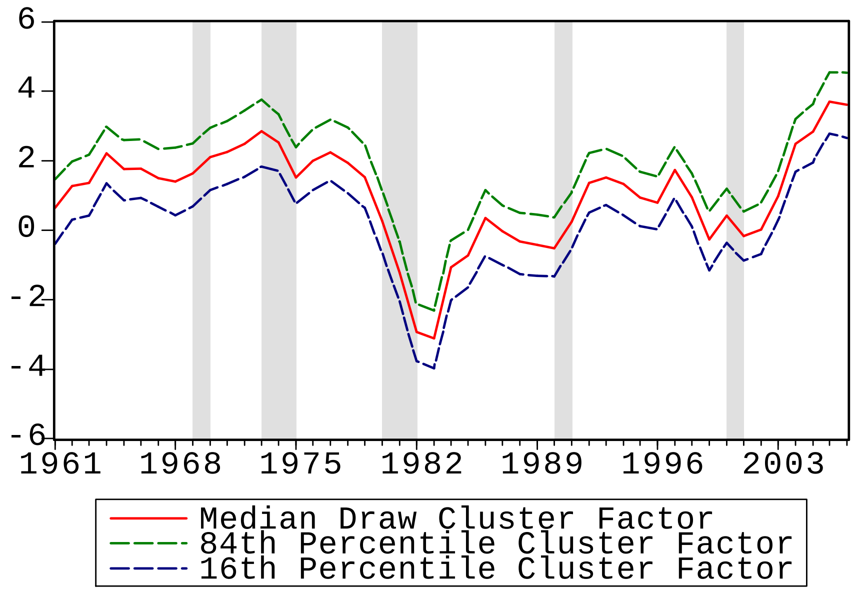

Figure 6: Cluster 3 Factor. Notes: The red solid line is the median of the posterior distribution of Cluster 3's factor. Dashed lines represent the 16th and 84th percentiles. Shaded regions are annual NBER recessions, where a recession is defined as a year in which any quarter was in recession according to the Business Cycle Dating Committee turning points.

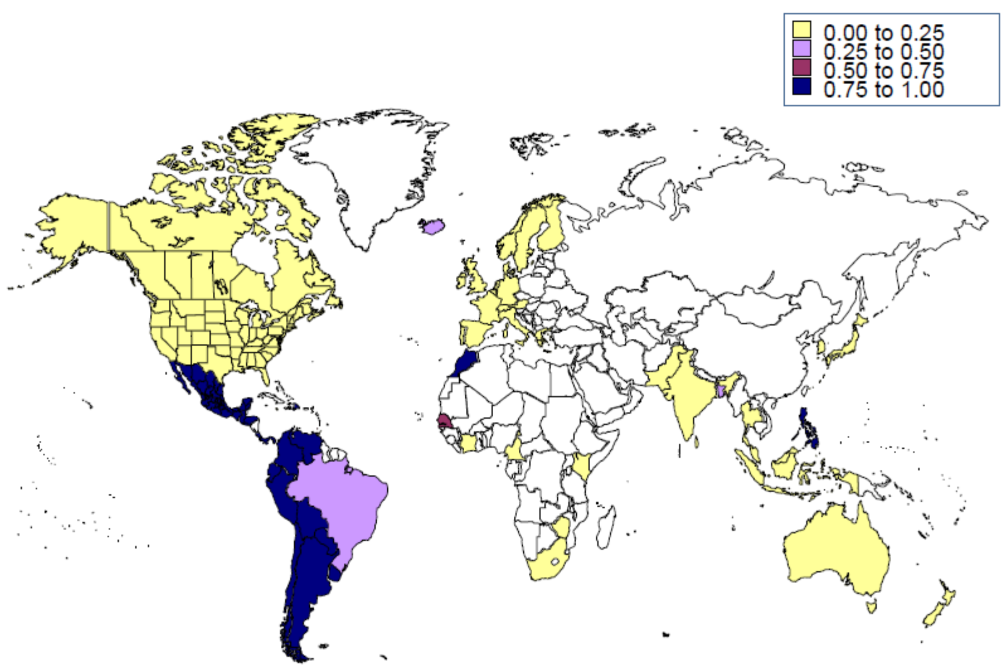

Figure 7: Cluster 3 Composition. Notes: The map shows the posterior probabilities of countries included in Cluster 3. Countries in white are omitted from the sample.

CHAPTER 3

THE DYNAMICS OF COMMODITY PRICES: A CLUSTERING APPROACH

In recent years the world has witnessed a commodity boom that has raised several questions and various explanations about the characteristics of commodity prices. Prices of grains such as corn, soybeans, wheat, and rice more than doubled during the $2006-2008$ peak period. Crude oil prices reached \$147 per barrel in July 2008 almost five times higher than what it was in 2003. The surge in prices created worldwide concern over energy costs and food security. The economic and social aspects of these price increases have led many researchers to speculate about the fundamentals of commodity prices. Commonly suggested determinants have been easy monetary policies, the devaluation of the dollar, excess liquidity, speculation in commodity markets, and high world demand.

The debates about what drives commodity price comovements had just commenced when the world was hit by the recent global downturn. With a sudden reversal around summer 2008, the soaring energy and food prices fell back to their 2006 values, signaling for a moment that this sudden price upsurge was nothing but a short-run phenomenon. But not long after, a second wave of rapidly increasing commodity prices came about. Beginning in May 2009, another surge in commodity prices is under way, reminding us what Frankel and Rose (2009) had pointed out: "...it cannot be a coincidence that almost all commodity prices rose together during much of the past decade..."

A rise in price of a single commodity would usually reflect something specific to that commodity and would not be informative about the overall economy. However, the synchronized movements of several commodities have different implications that a researcher should care about. These kinds of simultaneous movements could affect headline and core inflation of open economies or create concerns about food security for developing countries. Assume policy makers know which particular group of commodities share cycles and exhibit synchronized inflation and are also aware of the type of factors behind these dynamics. This would give them informational advantage in terms of which variables to carefully watch. If, say, oil prices, world demand, and interest rates are responsible for the upsurge in most of the commodity prices; then during expansionary periods when oil prices are trending up, contractionary monetary policy could dampen the price surge and prevent likely spillovers to core inflation.

The synchronicity of commodity prices is not new in macroeconomics. In a seminal paper, Pindyck and Rotemberg (1990) argue that there is "excess comovement" of seemingly unrelated commodities that cannot be explained by macroeconomic determinants such as interest rates or oil prices. They conclude that it is actually the herding behavior of market participants that causes the comovement in prices. Excessive or not, the consensus is that commodity prices do comove, whether it is through common macroeconomic fundamentals or through complementarity or substitutability in production and consumption or through a set of possible factors.

Most of the existing empirical work takes for granted that *all* commodity prices (or at least the ones defined under specific categories such as food and metals) move together (Baffes (2009), Hockman et al. (2010), Lombardi et al. (2010)). However, none consider how likely some group of commodities comove. If there are multiple factors driving primary commodities, different groups of commodities will share cycles due to different sources. We need to first identify the comovements of the commodities before we begin to talk about the determinants.

In light of above arguments, this paper addresses several questions. Which groups of commodities are likely to share cycles? Is there a common source behind the price comovements or are there multiple forces affecting different groups of commodities? How important are these possible factors or determinants behind the price dynamics and can we characterize them? To answer these questions we need to systematically decipher the correlation structure into its determinants, preferably with an empirical model suited for such an analysis.

The empirical model selection is important when it comes to analyzing the interrelations of many macroeconomic variables. Bernanke et al. (2005) suggest that VAR techniques suffer from a degrees of freedom problem, which puts restrictions on the number of variables that can be included in the system. They further emphasize the importance of dynamic factor models that can summarize the information from a large number of time series by a small set of indexes, or factors.[18] They propose a Factor Augmented VAR model to understand the common dynamics of many variables. In particular they apply a two-step approach: they first uncover the common space spanned by the factors of the data and then run a VAR of these estimated factors on possible determinants as a second step.

Likewise, in this paper I use a dynamic factor model to extract information from all the available non-energy commodity prices. I recognize that some comovements may not be simply due to a global factor like world demand but may also be a result of more sector-specific factors like droughts, floods or biofuel production that affect only smaller groups of commodities.[19] Novel to the paper is the use of an endogenous clustering procedure on a large data set of primary commodities to study these price dynamics. This approach was first introduced by

[18] Using dynamic factor models also provides advantages compared with the simpler cross correlation analysis that has been selected as a tool to investigate synchronous cycles.

[19] The estimation procedure applied in this paper requires that the determinants of the factors should be outside of the sample when conducting the factor model. Therefore, I focus only on non-energy products and use the crude oil prices in the ex-post analysis to determine their effects on the model factors and look for the validity of the claims made in the studies that list oil prices as the fundamental source of the price correlations.

Francis, Owyang, and Savascin (hereafter FOS, 2012) to study international business cycles.[20] Such an approach allows the data to freely choose from a set of possible unobservable factors and define their own groups. The empirical model will allow commodities to share similar cycles beyond that driven by a common global factor – an avenue overlooked by the literature. In particular, the empirical model includes a global factor and several group-specific factors where the groups are not defined a priori. After successfully extracting the fundamentals (factors) behind commodity price movements, I then try to characterize these underlying factors using Bayesian auxiliary least square regressions.[21]

The endogenous clustering analysis reveals that unrelated commodities that belong to the metal, agricultural materials and food families share cycles not only through global determinants but also through cluster factors. The global factor is the most important determinant for only vegetable oils (excluding olive oil), while the cluster factors carry greater importance for the rest of the commodities. Even though the commodity world seems to be coupling overall, there is still a considerable amounts of decoupling of particular commodities. In particular some commodity clusters show similarities in ways identified by specific product characteristics. For example, timber industry isolates itself from the rest of the agricultural raw materials and form a separate cluster. Likewise coffee forms another. Grains and vegetable oils decouple from the rest of the food category products and share most of their correlations through their cluster factor. Overall, commodity clusters found in this paper are not representative of standard narrow classifications (indexes) of commodity prices as defined by statistical agencies like International

[20] Factor models are extensively used to study various topics from international business cycles to regional analysis. Examples include: Kose et al. (2003, 2008), Hamilton and Owyang (2010), Neely and Rapach (2009).

[21] In a work similar to this, Vansteenkiste (2009) uses a clustered approach where she defines 4 groups exogenously for only 11 commodities. She tries only to link the global factor to possible macro factors while neglecting the cluster factors.

Financial Statistics (IFS). This implies that these narrower definitions, or subgroups, by these agencies do not consist of homogeneously moving products.

Bayesian regression analysis reveals that the world demand proxied by industrial production growth of many economies is a determinant of the global shock affecting all commodities. While vegetable oils and grain prices react to crude oil and fertilizer prices, other foodstuff are affected by the same variables as metals and materials. Through the cluster factor, metals, materials, and some food products seem to be reacting to a combination of many potential factors discussed in the literature: namely, interest rates, world demand, oil prices, and stock market indexes. Simple examination of these group-specific commodity comovements implies that in times of high oil prices and high growth in world production, low interest rates can amplify the increases in their price levels and further quantitative easing may indicate higher inflation in commodity prices which could make the economy more vulnerable to inflationary pressures.

The reminder of the paper is as follows: Section 2 provides motivation for the paper. Section 3 presents a review of the literature. Section 4 describes the empirical model and the estimation procedure. Section 5 describes the data. Section 6 reports the findings and finally section 7 concludes.

2.1 Motivation

Commodity prices carry great importance, with their potential impact on aggregate output and the balance of payments and transmission of business cycle disturbances across countries by connecting commodity exporters and importers from developed to developing countries (Borenzstein and Reinhart, 1994). Changes in commodity price levels can create inflationary pressures on an economy that could make monetary policies harder to conduct. If commodity production constitutes a larger percentage of aggregate output, then their price movements should be taken into account in the design of policy. The same is true even for the monetary authorities that target the core inflation rate, which excludes volatile food and energy prices. For instance, the Fed pays attention to and targets the core inflation rate, claiming that it has resulted in better forecasts than the headline inflation rate over the past 25 years.[22] It is true that the recent commodity price boom has not been reflected in core inflation largely because of the recent economic downturn, which resulted in strong disinflationary pressures, as the FOMC members expected. But what if the Fed is wrong about the expected moderation in global growth and high commodity prices *do* spillover the core inflation?

Commodities are used as inputs of production in many industries. For example, cotton is a major input for textile industry, which accounts for 4.6 percent of core personal consumption expenditure (PCE) inflation. Again, "Shelter" for the U.S. accounts for around 30 percent of core CPI and "Vehicles" around 6 percent. These groups (shelters, vehicles) include housing materials, equipment, and automobiles that are produced with extensive use of basic commodities such as copper, iron (used in steel production), rubber, timber, and lead. The price surge of

[22] Targeting core inflation has its own debate. Any central bank that wants to reconnect with households and businesses, which care more about food and energy price changes than bankers and hedge funds, should target headline inflation as suggested by Bullard (2011).

these commodities is expected to alter the cost structure of many industries and, hence, create high prices that can heat up an economy's inflation rates. That could make targeting inflation difficult and create an environment in which easy monetary policy could overheat inflation, like back in 1970's.

Identifying comovements of prices allows for diversification of inflationary risk not just for monetary authorities. For example, if economic agents in a commodity-exporting country were to know which commodities are likely to experience price increases and the degree to which commodities comove, then these agents (households and/or financial institutions) could diversify some of the risk by expanding the range of export commodities they invest in, sell, or hold. They could diversify by trading in commodities that have weak linkages and do not share common cycles with the commodities they currently export – a point stressed by Cashin et al. (1999). In financial markets, participants can settle their portfolio and investment decisions securely if information about comovements of commodities is known. Moreover, Lu and Neftci (2008) examine the use of commodity options to hedge against the vagaries of international commodity prices for developing nations.

Given the importance of commodity prices, any kind of theory that aims to investigate the policy implications of commodity price dynamics should rely on detailed empirical investigations. Without a diagnosis of the cause of price peaks, we cannot talk about policies that may alleviate the costs of price increases or take precautionary actions to prevent large fluctuations in prices that may result in a crisis.

Recent literature has looked for possible explanations of what has been driving the synchronized commodity price movements. Several factors are considered, from global factors such as high global demand to more market-specific factors such as the rise in biofuel production.[23]

[23] To cite a few: Krugman (2008), Wolf (2008), Frankel (2005, 2008), Calvo (2008), Lombardi et al. (2010), Baffes and Haniotis (2010), Lescaroux (2009), Cashin, McDermott and Scott (2002), and Vansteenkiste (2009).

The widely accepted view is that the correlations across commodities are solely a result of common factor(s) (Byrne et al. (2011), Vansteenkiste (2009), Cashin et al. (2002), Lescarux (2009)).[24] This may seem plausible at first, but it is incomplete at best. As Foerster et al. (2011) argue, additional cross correlations of any kind could contaminate the global factor, and if not taken into account, can lead to overestimation of the true nature and the contribution of common factors in explaining cross-product comovements. Using disaggregate industrial production data, Foerster et al. (2011) show that the common factors are contaminated by the unmodelled sectoral linkages. Likewise, common factors behind commodity price dynamics may reflect not only global shocks but also the propagation of idiosyncratic shocks within particular groups, usually by way of less pervasive factors.

The sparse factors can be thought of first as reflecting the different market properties across commodities. Shocks related to those specific markets may not spill over to other industries. In particular, shocks that generally emerge from climatic conditions or adverse weather such as floods and droughts directly affect agricultural products while their propagation into the mining industry is less likely. Similarly, for metals and even for some agricultural raw materials, such as timber or rubber, fertilizer costs may not be as relevant.

The traditional way to introduce the sparse industry factors is to exogenously group similar commodities; for example, one could ex-ante define "Food" and "Energy" clusters. While plausible, this may not be the best practice. Even within the same ex-ante categories we might have different underlying driving factors behind the commodity movements. In other words, assuming one "Food" sector will not allow for possible within-sector heterogeneity of

Hockman (2010) and Mueller (2011) look also at the biofuel effect on food commodities.

[24] Oil prices have been cited as the classic example of a common factor. As almost all industries are energy dependent (even when oil is not used directly in production, it is used in transportation), oil prices feed into the cost functions of almost all commodities.

particular commodities. Besides, seemingly unrelated commodities are argued to exhibit excess comovement.[25] Therefore if the ex-ante grouped commodities are "closely related", such groupings would not allow for unrelated commodities to share cycles other than through the global factor.

Due to the characteristics of the food commodities, natural disasters like floods and droughts might only affect some small group of products. Droughts in grain-producing regions over the last years have helped lower the world grain supply, which was thought to have significant impact on the grain price levels (Trostle, 2008). Australia has been suffering from a severe drought since 2004, which considerably reduced its production of agricultural products. Figure 8 presents the growth rate in total supply in metric tons for Australia during the drought period from 2004 to 2007. While total meat, vegetables, and corn supply showed big fluctuations, fish supply was relatively more stable compared with the other food commodities. One gets a similar graph for China, which has been experiencing the worst drought of their recent history. These observations suggest that drought may not have a significant effect on countries' seafood supply but severely influence the grains and livestock production. Therefore we may want to avoid grouping seafood with other grain and meat products.

Simple examination of the nature of commodities also advises against taking an ex-ante stance on commodity clusters. In particular, consider corn and rice. These two commodities experienced high prices during the price boom and their price rise was argued to be related to the same factor. Krugman (2008) argued that biofuel production caused farmers to expand the portion of their land used to cultivate more corn as it has become more profitable for farmers to invest the corn proceeds in ethanol production. This reduced the hectares of land used for

[25] Pindyck and Rotemberg (1994).

other grain plantings (e.g., soybeans). Since climatological and land conditions are different for corn and rice, farmers are unlikely to substitute land between them. While biofuel production might have had a direct effect on some grain products, we may not argue the same thing for rice crops.

From a more analytic view, FOS (2012) document the consequences of possible grouping (clustering) misspecifications in the Monte Carlo analysis they conduct. The idea of their simulation is to emphasize what may happen if the researcher unknowingly puts a time series in the wrong group and uses the traditional exogenously defined block factor approach in estimation.[26] They show that even small degrees of misspecification can cause a reduction in the model's overall fit. Specifically the entropy measure and mean square errors for the estimated factors increase with misspecifications, causing inconsistent model estimates.

Given these arguments I choose not to apply the traditional ways of defining groups of commodities; instead, I employ the FOS (2012) endogenously clustered dynamic factor model that places no initial restrictions on the groupings to study the synchronous movements of commodity prices. With the aid of this unrestricted model, the commodity groups can be formed based on any one or combination of any possible factors, such as those discussed above, without the fear of model misspecifications.

2.2 Literature Review

After a stable phase of commodity price inflation for over two decades, the late 2000's have seen price increases reaching record high levels and causing the world to experience one of the longest and broadest post World War II price booms. The previous price boom happened

[26] The time series in the simulation can represent many economic variables, such as a country's GDP or a commodity's price level, a city-specific housing price, or industrial production.

in the early1970's and was followed by a period of low levels in the 1980's. While commodity price levels maintained stability in the 1990's, nominal prices for grains (such as corn, soybean, palm oil, wheat, and rice), energy, and metals more than doubled during the 2006-2008 boom (see figure 2). The consequences for some developing nations were more severe than others. Riots and violent demonstrations over the soaring costs of basic food have been reported in many countries including Bangladesh, Haiti, Yemen, Egypt, Morocco, and Mexico. Due to the severe social aspect of high food and fuel prices, organizations around the world held meetings, and discussed possible coordinated policy actions and interventions in order to aid the societies that could not maintain sufficient dietary requirements. (Examples include the recent G-20 meeting, UNICEF Food Prices Increases/Nutrition Security: Action for Children and Food and Agriculture Organization's High-Level Conference on World Food Security.[27] Right after the crises, International Fund for Agricultural Development made available up to US$200 million to provide support for farmers). By the end of 2008, energy and food prices significantly declined in the wake of the financial crises and the global economic downturn. However, another surge in prices started in May 2009 and the rise continues as of the 3rd quarter of 2011.

All of the aforementioned changes in commodity prices raised interest in the determinants of such changes and a great deal of research has been devoted to understanding the comovements across commodities. Along the line of these studies, Calvo (2008) suggests excess liquidity and low interest rates as the cause of the recent price boom. Wolf (2008) blames it on increased world demand. Krugman (2008) argues that the increase in oil prices caused govern-

[27] G-20 meeting: Communiqué - Meeting of Finance Ministers and Central Bank Governors, Washington DC, 14-15 April 2011. Unicef link can be found at http://www.unicef.org/eapro/Food_Prices_Technical_Note_-july_4th.pdf. FAO Conference tried to make a strategy to deal with hunger and unrest resulted from soaring food and oil prices. Delegates of the conference also focused on increased biofuel production and how it relates to food production and prices. The conference however hit a snag over the debates about embargos and export restrictions.

ments to support biofuel production, which provides incentives for farmers to supply corn to be used in bio-ethanol production.[28] Farmers also switched land between corn and other grains, which reduced the overall supply for grains, which led to the increase in food prices.

In an attempt to summarize the studies about commodity price dynamics, Frankel and Rose (2009) list three competing theories explaining the recent boom. The first one is "global demand growth", which accelerated with the inclusion of high-demand countries such as China and India, causing the observed high prices.[29] Yet, this line of argument is criticized by researchers looking at the early effects of the sub-prime mortgage crisis that hit the U.S. in 2007. Economic growth downgraded for many countries lowering the production (hence the demand) for commodities globally during the time of crises, while commodity prices were still on the rise in the first 3 quarters of the recession, contradicting the obvious link between the two.

The second theory focuses on financial markets and argues that "speculation" was the main cause of the commodity boom.[30] Given there are futures markets for commodities, when market participants expect high prices they may hold long positions. If there is no particular reason to expect higher prices but the financial agents continue to do so, the resulting buying behavior can inflate a speculative bubble that creates high stock prices in commodity markets. Opponents of this explanation of speculative buying of commodity futures draw attention to the low inventory levels of commodities. As stated by Krugman (2008): if there were a bubble then we should have seen high inventories, which were not evident. However,

[28] Corn based ethanol is used to produce biofuels therefore ethanol production is usually used to proxy the biofuel production in empirical analysis.

[29] Wolf (2008), Svensson (2008).

[30] Citations include Hamilton (2009), Wolf (2008), Baffes and Haniotis (2010), Frankel (2008).

Frankel (2008) continues to acknowledge speculative explanations by claiming that inventories were not measured correctly. For example, the standard data exclude the amount of crude oil that still lies underground which is much larger than what has been measured as inventories.

The third and maybe the most popular explanation is “easy monetary policy." Low interest rates reduce the cost of holding inventories, since it is no longer profitable for the producers to invest the proceeds. Hence, by keeping interest rates at low levels, the Fed indirectly and presumably unwillingly causes decreased production and high prices. Furthermore, low interest rates create excess liquidity that can find its way into commodity markets as more and more people switch from Treasury funds to commodity contracts, thereby causing prices to rise.[31] Critiques of the interest rate channel use the same argument that was used against “speculation”: Where are the inventories?

Empirical investigations try to make theoretical links between commodity prices and their determinants. For instance, Baffes and Haniotis (2010) analyzed the effects of excess liquidity, speculation, food demand growth from emerging countries, and biofuel production on food prices. They found strong links between energy and non-energy commodities and less evidence for the effect of biofuel production on food prices. Instead they argue that it is the “new money,” the excess liquidity, which has found its way into the commodity markets and caused a speculative bubble – and hence the boom. Byrne et al. (2008) identify a common factor behind commodity price comovements by applying non-stationary panel analysis. Then they relate this common factor to potential macroeconomic variables using a FAVAR approach and find evidence of interest rate influence on commodity prices. Lombardi et al. (2010) run separate VARs for each of 15 non-energy products to look for effects of global industrial

[31] Frankel (2008), Wolf (2008), Akram (2009).

production, the U.S. effective exchange rate, the U.S. interest rate and the price of crude oil. They support the link between exchange rate and commodity prices and reject the effects of interest rates and oil prices.

Using dynamic factor analysis, Vansteenkiste (2009) investigates the relative importance of common factors for the non-fuel commodity price dynamics of 32 commodities for the period 1957-2008. She finds evidence of a common factor that becomes increasingly important throughout the sample period. As a robustness check, she also includes group-specific factors in estimation and looks for the effects of the global factor for 11 commodities.[32] She suggests that the global factor is more important than the group specific factors. However, her variance decomposition suggests this is true for only 3 (wheat, maize and cotton) out of the 11 products; for the rest the group specific factors seem more important. She later used IV regressions to test the potential effects of crude oil, fertilizer prices, dollar effective exchange rate, interest rate and global demand (proxied by the industrial production of OECD and 6 major non-OECD countries) on the extracted global factor component. She finds evidence that oil, exchange rate and interest rates are important. No attempt was made to characterize the group specific factors.

2.3 Empirical Model and Estimation Methodology

I employ a dynamic factor model where each commodity price inflation is affected by a global factor common to all and a block factor common to particular group to inflation series. Inflation rates across blocks can only be correlated through the global factor. There are in total

[32] The reason she reduced her sample for the exogenously defined factor analysis is that she only grouped the commodities she knows are in one way or anotherrelated. She avoids misspecification by this means. In this sense my analysis is the first one to introduce group-specific factors for the whole set of commodity price data without the fear of misspecification.

J block-specific factors (clusters) and one global factor. Assumptions of the dynamic factor model will be that the factors are unobservable and orthogonal to each other, and all cross-correlation of the series comes only through the factors, i.e., the variance-covariance matrix of the factors is diagonal. There are thus, K ($K = 1 + J$) dynamic factors to determine the comovements of inflation rates. As for the sector-specific factors, I follow the FOS (2012)'s endogenous clustering algorithm which gives the data the freedom to choose its own grouping.

Let I denote the number of goods and T denote the length of the time series. Then for an observable inflation rate denoted by $y_{i,t}$ for commodity i, we have

$$y_{i,t} = \alpha_i + \beta_{i,0} F_{0,t} + \sum_{j=1}^{J} \gamma_{i,j} \beta_{i,j} F_{j,t} + \varepsilon_{i,t}, \tag{9}$$

where α_i is a vector of intercepts; $\beta_{i,0}$ and $\beta_{i,j}$ are diagonal matrices of factor loadings; $F_{0,t}$ is the global factor affecting each series; F_{jt} are the group factors; and $\varepsilon_{i,t}$ is a series-specific idiosyncratic error term. As noted in FOS, $\gamma_{i,j} = \{0, 1\}$ is a grouping indicator that defines whether series i belongs to group j. Further, each series is restricted to one single group obtained by the restriction $\sum_j \gamma_{i,j} = 1$. Factor loadings are specific to each series, which allows for different responses of inflation in response to the same shock.

The evolution of each factor and the idiosyncratic error term are determined by an autoregressive equation of order q^f and q^ε, respectively;

$$F_{k,t} = \phi_{k,1} F_{k,t-1} + \phi_{k,2} F_{k,t-2} + ... + \phi_{k,p} F_{k,t-q^f} + e_{k,t} \quad for\ k \in \{1, ., K\}, \tag{10}$$

$$\varepsilon_{i,t} = \varphi_{i,1} \varepsilon_{i,t-1} + \varphi_{i,2} \varepsilon_{i,t-2} + ... + \varphi_{i,q} \varepsilon_{i,t-q^\varepsilon} + \epsilon_{i,t} \quad for\ i \in \{1, ., I\}, \tag{11}$$

where $e_{k,t}$ is a factor-specific idiosyncratic error term with variance λ_k^2, and $\epsilon_{i,t}$ is the idiosyncratic disturbance with variance σ_i^2. The disturbance terms, e and ϵ, are uncorrelated and each distributed normally with zero mean and their respective variances. As the factors are unobservable the sign of the factors and the sign of the factors' loadings have to be separately identified. Following FOS, I normalize the first element of each factor to be positive to overcome the issue. Another identification problem is to identify the scale of the factors. Here I follow KOW (2003), Sargent and Sims (1997), and others in assuming that λ_k^2 is constant.

I estimate the model presented in equation (9) with Bayesian Markov Chain Monte Carlo (MCMC) techniques. To sample the factors, I follow Kim and Nelson (1999) and apply Kalman filters. Once the factors are known (or given) I follow Chib and Greenberg (1995) to sample the model parameters. Sampling iteratively from the conditional distribution of the model's parameters given the factors and from the conditional distribution of the factors given the parameters is repeated many times. This is the essence of Gibbs sampling and under the regulatory assumptions (see Chip and Greenberg, 1995) these sequences of draws from the conditional distributions converge to the joint posterior density of the entire system, independent of initial values selected. The technical details of the estimation are provided in the appendix.

The endogenous clustering model represented here can be estimated using either Bayesian or classical techniques. The classical algorithm can solve the problem by forming every possible group and then employing a model selection criterion to determine the best clusters. However, with a large panel of data this grid search-like procedure would be inefficient and possibly infeasible.

The Bayesian approach offers several other advantages in estimation. First of all, Bayesian inference provides computational easiness for latent variable models like the one presented in this paper. As noted in Paap (2002) the likelihood function of classical estimations of these models includes many integrals and numerical integrations which make standard estimation models like maximum likelihood infeasible whereas the Bayesian MCMC approach only considers the likelihood function conditional on the simulated unobserved variables; therefore, it does not require computing the unconditional likelihood function of the model itself. This makes the estimation easier and much faster than most of the standard classical techniques. Paap also argues that Bayesian modeling allows a more convenient way of dealing with parameter uncertainty, which needs to be taken into account when dealing with unobserved variables.

Another important advantage of the Bayesian method concerns its small sample properties. It has been argued that MCMC computation works equally well for large and small samples.[33] Recently, with the wide use of the disaggregated data, researchers have utilized dynamic panel data econometrics. However, these models have been documented to perform poorly in estimation and inference without correcting for the small sample biases if the sample size is small. [34] The use of Bayesian methods offers an advantage in the sense that it does not require a correction when dealing with small samples. As Berger (1985, page 125) says Bayesian analysis would be preferable to any particular large sample classical techniques. He further adds that Bayesian analysis would be equivalent to the classical large sample procedures with large sample size and would be reasonable to use with moderate and small sample sizes where many classical estimation techniques fail.

[33] Western (1998), Martin (2005), Berger (1985).

[34] IMRR(2003), Chen and Engel (2005), Phillips and Sul (2007), and many others use several methods to account for small sample biases. The most commonly known small sample bias correction is Killian's bootstrap after bootstrap method. However this method has been proved to perform poorly with highly persistent series. The Andrews (1993) and Andrews and Chen (1994) median unbiased estimator is another way to correct for the bias. However this method does not work well if true AR(1) is near unity. Another method is by Pesaran Zhao (1999) who extends Killian's method for long-run coefficients that are nonlinearly dependent on the short-run ones. The Bayesian methodology performs equally well for large and small samples and provides an estimation tool that does not need any correction for small samples.

2.3.1 Model Selection

While cluster memberships are endogenously determined, the number of clusters still have to be exogenously selected. However we can also endogenize the number of clusters by computing the Bayes factors. This paper applies Chib (1995) for every Gibbs sample block and follows Chib and Jeliazkov (1995) where Metropolis Hastings is implemented. The basic marginal likelihood identity (BMI) of the model is given as:

$$\ln \widehat{m}\left(\mathbf{Y}\right) = \ln f\left(\mathbf{Y}|\mathbf{\Theta}^{*}\right) + \ln p\left(\mathbf{\Theta}^{*}\right) - \ln \widehat{p}\left(\mathbf{\Theta}^{*}|\mathbf{Y}\right), \tag{12}$$

where Θ is the full set of model parameters. The expression requires the evaluation of the log likelihood function $\ln \widehat{m}\left(\mathbf{Y}\right)$, the prior $\ln p\left(\mathbf{\Theta}^{*}\right)$, and an estimate of the posterior ordinate $\ln \widehat{p}\left(\mathbf{\Theta}^{*}|\mathbf{Y}\right)$ evaluated at a high density point $\mathbf{\Theta}^{*}$ (e.g., a modal point) of the posterior draws of the parameters.[35] The log likelihood and the priors at the modal values of the parameters can be directly computed from the whole set of posterior draws gathered from the Gibbs estimation of the model (the Gibbs output). However, the posterior ordinate needs to be estimated with additional Gibbs sampling steps of the same model but with reduced samplers. Details for sampling each posterior ordinate as well as the model likelihood and priors are all supplied in the appendix.

In this framework, the models are distinguished by the number of clusters. In particular, the empirical model and the BMI is estimated assuming a different number of clusters one at a time. Then the marginal likelihoods are used to decide which model to select. The model that maximizes the marginal likelihood reveals the optimum number of clusters. In principle

[35] Chib points out that the BMI holds for all values of Θ and the choice of the Θ^{*} is not critical but for the efficiency considerations Θ^{*} is selected to be a high density point so that the density can be accurately estimated.

one needs to calculate the Bayes factors of two models l and h, using the BMI as

$$B_{l,h} = \exp\left(\ln \widehat{m}\left(\mathbf{Y}\right)_l - \ln \widehat{m}\left(\mathbf{Y}\right)_h\right).$$

If $B_{l,h} > 1$, then model "l" is favorable to model "h". Comparing the bivariate Bayes factors for all models and finding the superior model is equivalent to maximizing the BMI given all models. Hence in the results section, I only list the BMI with its components.

2.3.2 Bayesian Linear Estimation on Factors

Once the factors are carefully extracted, in the next step I try to characterize them with additional analysis. Let Ξ_t represent the set of the variables we want to test on the factors (global and clusters), $F_{k,t}$. Then we can estimate the linear regression of the form

$$F_{k,t} = \varpi_k \Xi_t + \upsilon_k \tag{13}$$

where the error term υ is assumed to be normally distributed with mean zero and variance $\oplus_k^2$. The estimation procedure is a simple Gibbs application with two sampler blocks of parameters, namely ϖ and $\oplus^2$. The issue with such an estimation is that it makes use of the estimated factors as the regressand. The problems of using generated regressors are documented in Pagan (1984) and ideally one has to correct for the uncertainty coming from the generated regressor term for the inference to be correct as the posterior distribution of the model (13) depends on which factor $F_{k,t}$ is selected. .

If the factor is an observed variable, then the Gibbs application on (13) would converge to the unconditional posterior distribution of the parameters, i.e., $p(\varpi|\Xi)$. However if the factor is not observed and rather has a distribution, the Gibbs sampling yields a posterior distribution of the parameters conditional on the selected factor, say $p(\varpi|\Xi, \bar{F})$. Therefore to make inferences from $p(\varpi|\Xi)$, we can integrate $p(\varpi|\Xi, \bar{F})$ over the distribution of F :

$$p(\varpi|\Xi) = \int_F p(\varpi|\Xi, F)p(F|Y)dF, \tag{14}$$

where $p(F|Y)$ is the posterior distribution of the factor from the first step (factor analysis). Analytically we cannot solve this integral; instead we can approximate it by drawing large numbers of F from its posterior distribution and calculating the $p(\varpi|\Xi, F)$ by repeating the Gibbs sampling for each of these factor draws. This will result in an approximation of the unconditional posterior distribution of the parameters that we can make inferences from. Details are in the appendix.

2.4 Data

Monthly time series data for 42 non-energy commodity prices spanning from 1980 to 2011 are gathered from the International Monetary Fund (IMF)'s International Financial Statistics (IFS) Database. The commodities are selected on the basis of availability for the entire sample period. Fertilizer and energy prices are excluded so that they could be used in auxiliary regressions to see if energy prices are the main fundamental driving force behind the commodity price movements as argued in several studies. The details of the data can be found in the appendix.

Data are first seasonally adjusted (Census X12 multiplicative adjustment) and then converted to quarterly frequency mainly to increase the signal-to-noise ratio and to save some computational time. Since the empirical model requires stationary series, I log-difference the data, thereby computing the inflation rates for each commodity. Finally, I follow the factor model literature and normalize these inflation rates by demeaning each commodity price and dividing it by each series' standard deviation.[36]

Figure 10 plots the pairwise cross correlations of the commodity sample. Since simple cross correlations are static and cannot represent joint moves of many commodities these results should only serve as a preliminary check of possible linkages. The nominal prices of the product sample exhibits high positive cross correlations. Measures used in the literature such as concordance defines comovements as the same direction synchronized movements. However, the pairwise correlations suggest there are some products that exhibit negative relationships. Just because commodities move in the opposite direction does not necessarily mean that they cannot share the same source. The factor analysis presented in this paper do not exclude these kind of inverse movements and will recover commonalities, positive or negative.

The data for the auxiliary regressions come from several sources. The interest rate is proxied by the federal funds rate extracted from the Board of Governors of the Federal Reserve System. The exchange rate, Dow Jones stock market index, and U.S. house prices come from the St. Louis Fed's Federal Reserve Economic Data (FRED). The IFS database also provides crude oil and fertilizer prices (measured by phosphate rock). The federal funds rate and the U.S. effective exchange rate are deflated using U.S. consumer price index. The world demand is proxied by the industrial production of 30 countries from the IFS database where the countries are selected based on data availability. Ethanol production that accounts for biofuels is gathered from the Renewable Fuels Association. To measure climate changes, I use the global surface temperature anomalies from National Climatic Data Center (NCDC) of National Oceanic and Atmospheric Administration (NOAA) database.[37]

[36] There were some questions raised about standardization. I compared the results for both standardized and non-standardized data and found variance decomposition for both cases to be similar.

[37] As stated in the database, the anomalies are observed temperature departures from the 20th century (1901-2000) average of global temperature. An increase in these departures is evidence of global warming.

2.5 Empirical Results

2.5.1 The Optimum Number of Clusters via Bayesian Model Selection

The optimal model is the one that maximizes the marginal likelihood, which results in 4 clusters. Table 7 presents the details of the Bayesian Model selection outcome. Intuitively the posterior ordinates can be thought as a penalty of having additional clusters; and as one can see, the posterior ordinates are decreasing as the number of factor increases, thereby validating its purpose. The results also look consistent in the sense that the optimal model maximizes the marginal likelihood as well as the likelihood.

Table 8 presents in sample performance of the endogenously clustered factor model compared to the benchmark cases where the same data is estimated via a simple factor model with one global factor, a simple factor model with two global factors and an exogenously clustered factor model with clusters defined by IFS definitions.[38] The log marginal likelihood is maximized with the endogenously clustered model. Comparing marginal likelihoods in levels reveals that endogenously clustered factor model is 1.20×10^6 times more likely than the exogenously clustered model given the commodity price data used in the analysis.

2.5.2 Results for the Optimal Model

2.5.2.1 Inclusion Probabilities

Table 9 lists the commodities with their probability range across clusters. For each commodity the algorithm produces a posterior distribution of its indicator function. This means that each commodity has a probability, whether strong or weak, of belonging to each cluster.

[38] Based on the IMF industry classifications, the data can be clustered into 4 categories; Food, Agricultural Raw Materials, Beverages and Metals. Once I define these categories, I estimate the exogenously clustered factor model.

In table 9, I report the highest probability of belonging to one cluster for each commodity. Logs and wood inflation rates are strongly correlated through the first cluster. Lamb also belongs to this cluster with a weaker probability of 0.54.[39] However, since it constitutes only 20 percent of the cluster size and since it selects the cluster only half of the time, the corresponding factor should be dominated by the timber industry, reflecting the industry-specific properties of logs and wood.[40]

Consistent with observations in the commodity price boom, the second cluster consists of vegetable oils and grains. These products were responsible for the food price index spike more than any other commodities (Mueller et al., 2010), and it would contradict many related studies if they were not grouped together.

The third cluster consists of food, metal, and agricultural materials. The clustering analysis shows evidence that metals such as aluminium, copper, uranium, and zinc are strongly correlated with food products such as olive oil, fish, fishmeal, and sugar and weakly correlated with food products such as beef, lamb, oranges, bananas, and rice. Comovements of seemingly unrelated commodities are nothing new to the literature; for example, copper and wheat are found to be correlated by Pindyck and Rotemberg (1990). All metal prices in this cluster are highly correlated with other metals except iron. Cuddington and Jerrett (2008) offer empirical support for super cycles (long cycles for more than 15 years) of metal prices and posit the recent Chinese industrialization and urbanization as a likely cause. Therefore it would be interesting to check for this claim and test for Chinese demand on the cluster-3 factor.

Coffee products and iron belong to the last cluster. In particular, iron is the only

[39] The case of lamb is rather interesting since it also belongs to the cluster-3 almost equally likely (with a probability of 0.46).

[40] Lamb meat is 1 out of 5 commodities of cluster-1. Therefore it occupies 20 percent of the cluster size.

metal that does not share strong linkages with other metals. It is out of the scope of this paper to understand why iron shares cycles with coffee rather than with copper; however, this finding may open up an interesting avenue of research about these commodities. Moreover, the variance share of iron attributable to this cluster factor is only 4 percent (the complete variance decompositions listed under table 12); therefore, it would not be incorrect to claim that cluster-4 is defined by mostly the coffee industry and can be labeled as "Coffee Cluster".

Another interesting conclusion is the case of rice. Even though rice is listed under "grain products" in commodity price indices (along with barley, wheat, soybeans and corn) it does not share cycles with other grains – not even through the global factor.[41] What is more rice has the weakest probability of belonging to any cluster in the all the commodity sample used in this paper (it only achieves a maximum of 0.34 for cluster-3). What makes rice decouple and stand alone? This might stem from the fact that rice goes through different agricultural processes with specific needs for rainfall. Wheat needs a dry, mild climate to grow. Soybean fields should be well drained for its cultivation. And corn is a warm weather drought-resistant crop that requires relatively less moisture when developing toward maturity. Whereas, rice needs extreme humidity, and prolonged sunshine, it requires standing water throughout its growing period and is best suited for regions with high amounts of water supply. Other than these agricultural differences, country-specific effects (which are not accounted for in this analysis) may also be responsible for this "rice decoupling" since, apart from palm oil, cluster-2 products come from North America and United Kingdom while rice prices are taken from Thailand.

Overall, looking at the strongly correlated ($p(\gamma_{i,j}) > 0.9$) commodity cluster formations we can define 4 distinct categories (and I will refer to them as such hereafter): "Timber",

[41] The expained variance arrtibutable to the global factor for rice is 0. See the next section or appendix for the varaince-decompositions tables.

"Coffee", "Grains & Oils"(except olive oil) and a "Mixture" of agricultural raw products (e.g., wool, hides, rubber) metals (e.g., copper, lead) and food commodities (e.g., sugar, olive oil, salmon). This cluster formation provides evidence against distinct industrial categorization of commodities. In other words, these clusters are not representative of standard narrow classifications (indexes) of commodity prices as defined by statistical agencies. In particular, food products are spread across all clusters (weather weakly or strongly): While some of them share cycles with metals and agricultural raw materials, some of them decouple from the rest of the food sector (coffee and rice).

In related work, Vansteenkiste (2011) sets up an exogenously defined clustered factor model. She pools jointly produced or consumed commodities together into groups and defines 4 clusters: (1) Coffee and cocoa, (2) cotton-maize-sugar-wheat, (3) palm oil and soybean oil, and (4) copper-zinc-lead. This paper provides evidence against this kind of cluster formation. In particular, once the data are free to form their own clusters, cocoa and coffee fall into different groupings and sugar does not find its way into the same cluster as grain commodities.

2.5.2.2 Variance Decompositions

This section reports the variance decompositions where the clusters are constructed with the posterior values of the indicator function, $\gamma_{i,j}$. Since an observation may change clusters over the Gibbs iterations, we need a fixed estimate for the indicator function for the variance calculations. Table 10 reports the "weak probability variance decompositions," where an observation is assumed to belong to a particular cluster if it picks that cluster for the majority of the Gibbs run, i.e., $i \in j \; if \; p(\gamma_{i,j} = 1) > p(\gamma_{i,k} = 1) \; \forall k$. In this case all the commodities are matched to one cluster and we get the cluster memberships exactly as listed in table 9.[42]

[42] One of the drawbacks of such a calculation is that some of the interrelations among commodities are rather weak. Weakly correlated products share smaller portion of their cycles through factors which reduce the

Looking at the average variance decomposition percentages of table 10, the global factor is not playing a significant role in explaining the bulk of the commodity sample developments except for cluster-2 ("Grains & Oils). The cluster factors on average explain about 27 percent of commodity price variations, which dominates the effect of the global factor, which is only 7.2 percent. In particular, for cluster-1 ("Timber") and cluster-4 ("Coffee") the global factor is negligible; the market and production processes of these products may be too specific and closed to global developments. Overall Table 10 suggests that the dominant source behind commodity price comovements is the interrelations that come through more sparse cluster factors. This finding contradicts many studies that assume only one or two commons factors behind commodity price dynamics.[43]

The simultaneous move in prices of grains (corn, soybeans, wheat, rice) oils and meat have led many studies to agree on the existence of a single shared source for the food sector (Byrne et al. 2010; Vansteenkiste, 2009). To have a better insight into this claim, Table 11 highlights the variance decompositions of these products. Surprisingly, the source of the fluctuations seems to be different even for similar grain products that are in the very same cluster-2 ("Grains & Oil") such as corn and wheat. Corn, soybeans, and soybean meal are highly dominated by the cluster factors, while vegetable oils are mainly driven by the global factor. However for wheat, rice, and meat, it is idiosyncratic shocks that matter the most.[44]

In summary, the world factor does not seem to have a strong effect on corn, rice, wheat, meat, soybeans, and soybean meals prices, which invalidates explanations such as those that

variance decompotion values for each cluster. Once the commodities that have probability of belonging to a cluster below 0.9 are excluded from the data sample, the effects of global and cluster factors get increasingly large, explaining 46 percent of the whole sample variations compared with 34 percent of "weak probaility variance decomposition".

[43] Byrne et al. (2010), Vansteenkiste (2010), Cashin et al. (2010)), Lombardi et al. (2010).

[44] Rice and meat belong to cluster-3 decoupling from the rest of the food commodities investigated in table 12.

assert food commodities move together mainly due to, for example, increased world demand. Rice in particular is dominated by market-specific factors rather than aggregate factors as discussed in the previous section. The global factor does not have a significant effect on rice price fluctuations, which contradicts Vansteenkiste's (2009) findings where she finds 12 percent of rice price variations resulting from the global factor.[45]

Overall, the premise is that there is no single common factor driving the major commodity prices over time; instead commodities are interrelated through a set of cluster factors which contribute to the recent price peak more than the common factor. "Timber" and "Coffee" decouple from the rest of the sample, exhibiting different and probably more product-specific sources. However, more in-depth analysis is needed on the global and cluster factors to validate such claims.

2.5.2.3 Characterizing Factors

Figures 11 to 15 plot the factors along with NBER recessions dates.[46] The downturns in the commodity factors coincides with the U.S. recessions. In particular the global, second, and third cluster factors show a great slump during the Great Recession of the late 2000's, which suggests all of them were affecting the price fluctuations of food, metals, and materials during the commodity price burst. The big fall that corresponds to year 1994 in cluster-4 ("Coffee",

[45] The difference is likely to originate from empirical model specifications. She uses a dynamic factor approach with one global factor and her global factor is suffering from overestimation due to the additional correlation of commodities that are not accounted for in her analysis.

[46] The factor loadings are almost all positive for commodity prices expect a few commodities. Namely, swine which belongs to cluster-4 is negatively related to the cluster-4 factor and hard logs, shrimp, soft Logs, soft sawnwood, soybean meal and again swine are negatively related to the global factor. Looking at the individual variance decompositions (see appendix) the average explained variation of these commodities attributable to the global factor is only 1.85. Which implies that the average inflation for those commodities due to the global factor during recession is negligible. Besides, once combined with the cluster factor effect (which explains 35 percent of their variations) the overall impact will be positive.

figure 8) is consistent with the Brazilian coffee plantation expansion and Vietnam's entry into the market, which put pressures on the supply and lowered coffee prices.[47]

The variance decompositions of the previous section provided some evidence of multiple important factors behind the commodity price comovements. This section focuses on identifying these factors to see if any macroeconomic variables are correlated with the estimated factors.

In order to highlight the sources of these factors, I run Bayesian auxiliary regressions of the estimated factors on possible determinants that have been mentioned, argued, or strongly supported in the literature. In a related paper, Bryne et al. (2011) use a two-step FAVAR approach as described in Bernanke et al. (2005) and relate the common factor to the real U.S. short-run interest rate, global demand as proxied by U.S. real GDP growth, real crude oil prices, and risk measured by standard deviations of closing value of Dow Jones average. Vansteenskiste (2009) also employs a dynamic factor approach and test the global factor on possible determinants. In particular I use variables similar to those of Vansteenkiste (2009) – namely; the federal funds rate, U.S. dollar effective exchange rate, fertilizer prices, industrial production, and stock market index. In addition to these variables, I also test for U.S. housing prices, biofuel production, Chinese demand, and climate changes. Detailed description of the variables used in this paper are listed below.

1. Deflated Effective Federal Funds Rate (FFR), Quarterly: The nominal rate is deflated by the Consumer price index. Given the arguments in the literature, it is expected to have a negative impact on the commodity prices.

2. U.S. dollar Real Effective Exchange Rate (EER), Quarterly: Devaluated dollar (represented as a fall in the EER) causes the commodities to get cheaper in terms of foreign

[47] The "Timber" cluster factor (figure 4) seem to be recovered from big fluctuations of the 80's and early 90's and is relatively stable during the rest of the 2000's.

currencies, which in turn puts on a positive pressure on the prices.

3. Dow Jones Stock Market Index, Quarterly: It is used to measure the speculation bubble effects on the commodity price. It should result in higher commodity prices implying an expected positive sign in the regressions.

4. World Industrial Production, Quarterly: I use the quarterly industrial production of 30 countries, which includes developed economies as well as emerging and underdeveloped countries to proxy for world demand.[48]

5. Crude oil Prices, Dubai, Quarterly: Oil price increases cause a cost increase and higher commodity prices.

6. Fertilizer Prices, Quarterly: This should result in higher prices for many food and some agricultural materials (such as wool) due to cost increases.

7. Housing Prices, Quarterly: Recent subprime mortgage crises have spread around the globe and initiated the latest Great Recession. The burst in housing prices may not directly have caused the commodity price boom; however, it may have reduced the demand for its basic inputs: logs, metals, and materials. Hence, we can expect to see a positive relationship between house prices and their input prices.

8. Ethanol Production, Annual: To account for the increase in biofuel production, I use its main ingredient – corn-based ethanol production. High ethanol production growth could cause high food prices, especially for grains and oils, due to reasons described previously in this paper.

[48] The countries include Australia, Austria, Barbados, Belgium, Canada, Denmark Finland, France, Germany, Hungary, India, Ireland, Israel, Italy, Japan, Jordan, Republic of Korea, Luxembourg, Malaysia, Mexico, Netherlands, Norway, Portugal, Senegal, Spain, Sweden, Switzerland, Turkey, United Kingdom, and United States.

9. China Volume of Exports and Imports (China Trade), Quarterly: Emerging countries, especially China, took on a larger role in world trade while increasing the demand for commodities as well as their prices. The widely used measures for a country's demand are its industrial production or its gross domestic product. However, both of these variables for China are not available in quarterly frequency in 1980-2011 time period. Therefore, as an alternative, I use the total volume of exports and imports to account for quarterly Chinese demad.

10. Chinese Gross Domestic Product (China GDP), Annual: As discussed above, this variable is included to account for the increased Chinese demand. The cross correlation between volume of trade and GDP is 0.95; therefore, I do not use these two variables in the same regression to avoid multicollinearity issues.

11. Weather Anomalies, Quarterly and Annual: Climate changes could cause adverse weather conditions that could affect crop growth and reduce agricultural supplies. These anomalies are provided as departures from the 20th century average (1901-2000) and can be used as measures of adverse weather. An increase in temperature anomaly is an harbinger of global warming, which indicates the possibility of adverse weather reactions. To construct quarterly data, I aggregate monthly values for these global temperature anomalies across seasons.

The estimated factors are measured at a quarterly frequency. As a result, I conduct two separate analyses. First I look for the contemporaneous relationship of the quarterly factors with the interest rate, exchange rate, Dow Jones stock market index, world industrial production, U.S. housing prices, fertilizer prices, and climate anomalies. Additionally, to test for the biofuel effect (which is not available in quarterly frequency), I estimate annual regressions using annualized factors along with annual variates of everything listed above while substituting China Trade with China GDP. The caveat of aggregating the factors is that the regression results can potentially lose short-run information and can suffer from aggregation bias.

Quarterly Regression Results: Table 13 shows the results where each column represents a separate regression of the determinants listed in the rows. The global factor looks like it is capturing the world industrial production. The intuition is straightforward. When industrial production increases, demand for metals and materials accelerates. Higher income due to higher production tends to increase the demand for food, thereby spreading around the effects of the high global demand to almost all sectors.

Chinese volume of trade seems to have a significant effect on the factor that drives the correlation structure of the cluster-1 ("Timber"), which includes wood and logs from two countries: U.S. and Malaysia. So how can we link Chinese trade with these commodities? China is one of the top five importers of Malaysian timber. Chinese buyers also turned to the U.S. and Canada for wood after 2007 when Russia imposed higher tariffs on its logs. Chinese lumber imports from North America more than quadrupled from 4 percent in 2005 to 18 percent in 2010.[49] This revived the U.S. timber industry back from a depressed state since the subprime mortgage crisis. The Wallstreet Journal reports 10 to 15 percent expected increase in log harvests from big U.S. timber companies due to the recent export surge from China. These may be the reasons why we see a significant Chinese demand on cluster-1 ("Timber") factor.

Crude oil prices are found to be significant for cluster-2 ("Grains & Oils"). The farming sector is highly energy intensive; therefore, oil prices affect its cost structure. For example,

[49] Source: International Centre for Trade and Sustainable Development Bridges Trade BioRes, Volume 11, Number 18, 17th October 2011, pp 14.

Murray (2005) draws attention to the high use of fossil fuels in the U.S. farming industry with this comparison: "The U.S. food system uses over 10 quadrillion Btu ($10,551$ quadrillion Joules) of energy each year, as much as France's total annual energy consumption." Growing food without packaging, storage, or transportation accounts for 20 percent of this total amount. Besides, food travels from farms to distributors around the world, which amplifies the industry's dependence on fuel use. The conclusion is simple: fossil fuel reliance can alter grain commodity prices and hence can be reflected as this cluster's factor.

Most of the variables discussed in the literature as potential determinants of the commodity prices; namely, federal funds rate, speculation, world demand, and crude oil prices are found to be a part of "Mixed" cluster commodity commovements. Mining and manufacturing metals are energy intensive which help to link oil prices to this cluster-3 factor. From a monetary policy perspective, the federal funds rate can affect the dynamics of a large group of commodities of metals, materials, and some foodstuff. Intuitively this means that if the Fed keeps its quantitative easing policies in effect at times of high oil prices and high world demand, it would amplify the increases in commodity price inflation.

Lastly, cluster-4 ("Coffee") factor fails to highlight a significant presence of any of the variables tested in this analysis. This could be due to missing important macro variables or simply because this cluster represents shocks specific to the coffee industry that cannot be accounted for easily. For example, coffee markets are highly controlled by the International Coffee Organization (ICO), which was formed in 1963 in an attempt to stabilize prices through international cooperation. With members that account for 97 percent of world coffee exports and 80 percent of world coffee imports, ICO claims to achieve a balanced and sustainable world coffee economy and promotes coffee consumption. For example, it launched CoffeeClub Network in 2008 and implemented the Coffee Quality Improvement Programme in 2002 in order to stimulate demand through better standards of quality. The effects of these acts and agreements are likely to have an impact on the coffee industry, but are hard to measure.

Annual Regression Results: This section adds the remaining regressors – namely, ethanol production and China demand as measured by GDP – to the regression analysis and lists the findings in Table 14. The significant variables in each regression are consistent with the findings from the quarterly regressions. Additionally, speculation, fertilizer prices, and the U.S. dollar-effective exchange rates become significant to clusters 1, 2, and 3, respectively.

The global factor helps to feed the effect of global demand into the commodity prices as also shown in quarterly analysis. Cluster-1 ("Timber") factor now adds speculation to its possible determinants. Cluster-2 ("Grains & Oils") is affected by fertilizer prices, which is not a surprising result as fertilizers are used in cultivation to improve plant growth. In particular, corn, soybeans, and wheat are the three major crops associated with high consumption of fertilizers.

Cluster-3 "Mixed" factor is now explained by the changes of the exchange rate along with the federal funds rate. The devaluation of the dollar may reduce the price competitiveness of non-U.S. countries and diminish production of exporting goods for these countries. On the demand side, a devalued dollar can cause prices of a product to become cheaper in foreign currencies, this may increase the demand and alter the price dynamics. The combination of both supply and demand effects can accelerate the increase in its price level.

Finally, world demand affects the last cluster ("Coffee"). This could be linked to successful attempts of the ICO's coffee demand stimulation acts described previously.

Looking for the impact of biofuels on commodity prices, I cannot provide evidence in favor of Krugman's (2008) argument that increased biofuel production is one of the main causes of the grain price surge.

2.6 Conclusion

The dynamics of commodity prices have been changing over the last half of the decade. No economy is immune to inflation, and if price increases are synchronized and remain persistent enough they can pass through to the core inflation rate, creating a need for action by the monetary authorities. This paper investigates the comovements of commodity prices and what drives them from a statistical point of view. Summarizing information from a large panel set of commodity prices, I find that commodity cluster compositions do not represent the standard narrow classifications (indexes) of commodity prices as defined by statistical agencies like International Financial Statistics (IFS). For example, timber products isolate itself from other agricultural raw materials, and form a separate cluster. Likewise coffee forms another. I also find another cluster of commodities consisting of seemingly unrelated products, such as metals, agricultural materials, and some food products. Additional analysis to characterize these correlations indicates the importance of the federal funds rate, high world demand, high crude oil prices, fertilizer prices, Chinese demand, and speculation in financial markets in driving these products' common movements.

2.7 Tables and Figures

BMI Estimation-Model Selection

No.of Clusters	$\ln f(\mathbf{Y}\|\mathbf{\Theta}^*)$	$\ln p(\mathbf{\Theta}^*)$	$\ln\widehat{p}(\mathbf{\Theta}^*\|\mathbf{Y})$	$\ln\widehat{m}(\mathbf{Y})$
2	-6123	-1036	308.6	-7468
3	-5954	-1048	278.5	-7287
4	**-5797**	-1051	269.5	**-7118**
5	-5802	-1056	278.0	-7137
6	-7689	-1068	222.5	-8979

Table 7: BMI Estimation - Model Selection. Notes: First column shows the likelihood of the model, Second column refers to the prior value at the modal points. Third column represents the sum of the posterior ordinates calculated from the reduced runs described in the appendix. And finally last column shows the model marginal likelihood.

Model Comparison

Model	log ML	Diff.	Ratio
Basic Factor Model (1 global)	-9211.4	-2092.6	inf
Basic Factor Model (2 global)	-7627.9	-509.1	1.25×10^{22}
Exo. Block Factor - IMF definitions	-7132.8	-14	1.20×10^{6}
Endo. Block Factors	$\mathbf{-7118.8}$	n/a	n/a

Table 8: Model Comparison. Notes: The table shows the log marginal likelihood for different factor models. The first row presents the outcome from a basic factor model with one global factor and no cluster factors. The second model has two global factors and no cluster factors. The third one represents the exogenously clustered factor model with IMF definitions for clusters (Food, Metals, Beverages and Agricultural Raw Materials). The last model shows the results from endogenously clustered factor framework. All models are estimated with the same commodity price data. The first column shows the logarithm of the marginal likelihoods. The second column shows the difference in the log marginal likelihoods between the best model and each other model. The last column shows how much more likely the best model is compared to each other model.

Inclusion Probabilities Across Clusters

	Cluster-1	Cluster-2	Cluster-3		Cluster-4
$p(\gamma) \geq 0.9$	Hard Logs	Barley	Aluminum	Tin	Coffee, Arabia
	Hard Sawnwood	Corn	Copper	Wool,Coarse	Coffee, Robusta
	Soft Logs	Soybean Oil	Fishmeal	Wool,Fine	
	Soft Sawnwood	Wheat	Hides	Zinc	
		Soybean Meal	Lead	Salmon	
		Soybeans	Nickel	Sugar,US	
		Palm oil	Olive Oil	Uranium	
		Canola oil	Rubber	Sugar,World	
$0.8 \leq p(\gamma) < 0.9$		Cotton	Cocoa		
		Groundnuts			
		Poultry			
		Sunflower Oil			
$0.5 \leq p(\gamma) < 0.7$	Lamb				Iron
$0.3 \leq p(\gamma) < 0.5$			Beef		Shrimp
			Oranges		Swine
			Tea		
			Bananas		
			Rice		

Table 9: Inclusion Probabilities Across Clusters. Notes: The table summarizes the posterior inclusion probabilities for each cluster and lists the members. For each commodity highest probability of belonging to one cluster is reported.

Variance Decompositions Across Clusters

Factor	Cluster 1	Cluster 2	Cluster3	Cluster4	Sample Average
Global	1.04	18.69	3.64	0.86	7.2
Cluster	37.74	31.29	20.09	35.36	26.9
Global +Cluster	38.78	49.98	36.22	23.74	34.1
Idiosyncratic	61.2	49.98	76.24	63.8	65.9

Table 10: Variance Decompositions Across Clusters. Notes: The table summarizes the variance decompostion in percentages where the clusters are estimated with the endogenous clustering algorthim.The clusters are constucted with the posterior values of the indicatior function. An observation is assummed to belong to one cluster if the said observation picked that cluster more of the time over the Gibbs run than the other clusters.

Product Name	Global	Group	Idiosyncratic
Corn (2)	8.8	**46**	45.2
Soybeans (2)	2.3	**92.8**	4.9
Soybean meal (2)	5.3	**90.3**	4.3
Soybean oil (2)	**49.3**	39.1	11.6
palm oil (2)	**49.4**	14	36.6
canola oil (2)	**52.6**	11.9	35.5
Wheat (2)	4.2	21.9	**73.9**
Rice (3)	0	7.7	**92.2**
Meat (3)	6.5	1.7	**91.8**

Table 11: Variance Decompositions for Grains, Oils and Meat. Notes: The table summarizes the variance decompostion in percentages. Members are allocated to clusters that they picked the most over the Gibbs run. The paranthesis indicates each commodity's selected cluster. Bold values represents the highest variance decomposition for each product.

Variance Decompositions for All Commodities

Product Name	Global	Group	Idio.	Product Name	Global	Group	Idio.
Hard Logs (1)	1.2	72.6	26.2	Copper (3)	5.9	60.7	33.3
Hard Sawnwood (1)	0	68.8	32.2	Fish-Salmon (3)	2.1	11.2	86.7
Lamb (1)	0.4	10.1	89.5	Fishmeal (3)	1.1	18.8	80.1
Soft Logs (1)	2.9	25.4	71.6	Hides (3)	0.8	16.1	83.3
Soft Sawnwood (1)	0.7	11.8	87.5	Lead (3)	1.9	22.1	76.1
Barley (2)	14.8	31.1	54.1	Nickel (3)	3.4	40.9	55.7
Corn (2)	8.8	46	45.2	Olive Oil (3)	0	16.6	83.3
Soybeans (2)	2.3	92.8	4.9	Oranges (3)	0.3	1.4	98.3
Soybean meal (2)	5.3	90.3	4.3	Rubber (3)	9.5	45	45.5
Soybean oil (2)	49.3	39.1	11.6	Sugar, Free Market (3)	0.7	7.8	91.5
palm oil (2)	49.4	14	36.6	Sugar, US (3)	1.2	2.9	95.8
canola oil (2)	52.6	11.9	35.5	Tea (3)	0.8	2.9	96.2
Wheat (2)	4.2	21.9	73.9	Tin (3)	13.5	15.5	71
Cotton (2)	6.5	14.1	79.3	Uranium (3)	1.9	12.6	85.6
Groundnuts (2)	1.6	5.5	92.8	Wool, coarse (3)	7,2	23.9	68.8
Poultry (2)	0.1	4	95.9	Wool, fine (3)	9.5	23.5	67
Sunflower Oil (2)	29.4	5.6	64.9	Zinc (3)	2.6	41.6	55.8
Rice (3)	0	7.7	92.2	Coffee, Robusta (4)	0.7	80.9	13.4
Meat (3)	6.5	1.7	91.8	Coffee, Other (4)	0.2	85.5	14.3
Aluminium (3)	6	55.8	38.2	Iron (4)	2.4	4	93.6
Cocoa beans (3)	4.9	12.6	82.5	Shrimp (4)	0.7	1.8	97.5
Bananas (3)	0.4	0.7	98.9	Swine (4)	0.3	4.6	95.2

Table 12: Variance Decompositions for All Commodities. Notes: The table summarizes the variance decompostion in percentages. The members are allocated to clusters given the modal value of the cluster probability. The paranthesis indicates each commodity's selected cluster.

Quarterly Regression Results

Varible Name	Global	Cluster-1	Cluster-2	Cluster-3	Cluster-4
Federal Funds Rate	-0.02	-0.01	-0.02	-0.05*	-0.02
	(-0.06 0.02)	(-0.05 0.03)	(-0.05 0.02)	(-0.08 -0.02)	(-0.05 0.02)
World IP	0.22*	-0.16	0.05	0.34*	0.18
	(0.04 0.39)	(-0.35 0.03)	(-0.13 0.23)	(0.18 0.51)	(-0.02 0.38)
Dow	-0.02	-0.002	0.01	0.04*	-0.02
	(-0.04 0.01)	(-0.03 0.03)	(-0.02 0.04)	(0.01 0.06)	(-0.05 0.01)
Oil Price	0.01	0.01	0.01*	0.02*	-0.002
	(-0.00 0.02)	(-0.00 0.02)	(0.008 0.03)	(0.01 0.03)	(-0.02 0.01)
Fertilizer Prices	0.01	0.01	0.005	-0.01	0.002
	(-0.02 0.00)	(-0.01 0.02)	(-0.01 0.02)	(-0.02 0.002)	(-0.01 0.02)
US House Price	-0.06	0.06	-0.001	0.02	-0.16
	(-0.22 0.09)	(-0.10 0.24)	(-0.16 0.16)	(-0.11 0.16)	(-0.34 0.02)
Exchange Rate	0.001	-0.001	0.000	-0.001	0.002
	(-0.01 0.01)	(-0.01 0.01)	(-0.01 0.01)	(-0.01 0.00)	(-0.00 0.01)
China Trade	0.000	0.04*	-0.004	-0.001	0.02
	(-0.02 0.02)	(0.02 0.07)	(-0.02 0.02)	(-0.02 0.02)	(-0.01 0.04)
Climate Anomaly	0.35	0.19	0.25	-0.12	0.01
	(-0.28 0.98)	(-0.49 0.87)	(-0.39 0.89)	(-0.67 0.42)	(-0.64 0.66)

* denotes statistical significance

Table 13: Quarterly Regression Results. Notes: Each column represents a seperate Bayesian Regression on the variables listed in rows. Constant is excluded as in estimation. China Trade is measured as the volume of exports and imports. Variables except FFR and Exchange Rate are all percentage growth rates. FFR and Exchage rate are in deflated levels. Credible Intervals that are measured by the 5th and 95th percentiles are shown below of each coefficient.

Annual Regression Results

Variable Name	Global	Cluster-1	Cluster-2	Cluster-3	Cluster-4
Federal Funds Rate	-0.001	-0.037	-0.009	-0.06*	-0.04
	(-0.06 0.05)	(-0.02 0.09)	(-0.05 0.036)	(-0.11 -0.01)	(0.1 -0.01)
World IP	0.14*	-0.098	-0.074	0.037	0.13*
	(0.25 0.3)	(-0.02 0.22)	(-0.16 0.02)	(-0.06 0.136)	(0.02 0.24)
Dow	-0.02	0.02*	0.002	0.01	-0.01
	(-0.04 -0.002)	(0.001 0.04)	(-0.01 0.02)	(-0.01 0.03)	(-0.03 0.01)
Oil Price	-0.001	0.01	0.002	0.0073	-0.0085
	(-0.01 0.01)	(-0.001 0.02)	(-0.01 0.01)	(-0.00 0.02)	(-0.02 0.001)
Fertilizer Prices	0.003	-0.001	0.01*	-0.003	-0.001
	(-0.01 0.01)	(-0.01 0.01)	(0.00 0.01)	(-0.01 0.00)	(-0.01 0.01)
US House Price	-0.017	0.044	0.028	0.059	-0.008
	(-0.09 0.06)	(-0.04 0.13)	(-0.03 0.09)	(-0.01 0.12)	(-0.08 0.07)
Exchange Rate	-0.004	-0.011	-0.004	-0.01*	-0.01
	(-0.01 0.01)	(-0.02 0.001)	(-0.01 0.01)	(-0.02 -0.002)	(-0.02 0.00)
China GDP	0.0154	0.04*	0.026	0.024	0.03
	(-0.02 0.05)	(0.003 0.07)	(-0.00 0.05)	(-0.01 0.05)	(-0.06 0.003)
Bio Fuel	-0.001	0.003	-0.009	0.01	0.016
	(-0.02 0.02)	(-0.02 0.02)	(-0.02 0.004)	(-0.00 0.03)	(-0.001 0.03)
Climate Anomaly	-0.21	0.75	-0.48	-0.19	-0.43
	(-0.97 1.4)	(-0.5 1.96)	(-1.52 0.57)	(-1.3 0.92)	(-1.65 0.78)

* denotes statistical significance

Table 14: Annual Regression Results. Notes:Each column represents a seperate Bayesian Regression on the variables listed in rows. Constant is excluded as in estimation. Variables except FFR and Exchange Rate are all percentage growth rates. FFR and Exchage rate are in deflated levels. Credible Intervals that are measured by the 5th and 95th percentiles are shown below of each coefficient.

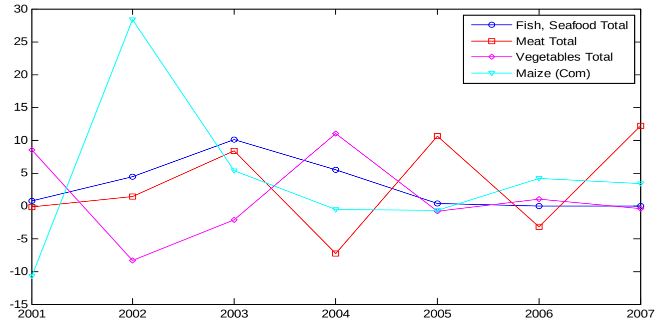

Figure 8: Supply Growth. Notes: Total supply growth in metric tones of Maize, Meat (total), Seafood (total) and Vegetables (total) for Australia during the draught period. Annual data is gethered from Food and Agriculture Organization of the United Nations Statistics (FAOSTAT).

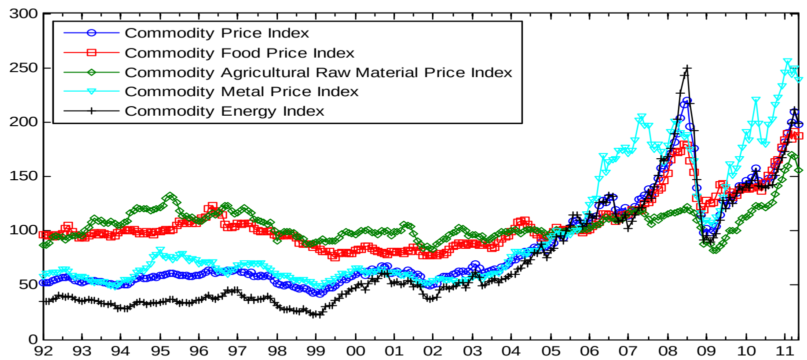

Figure 9: Commodity Price Indexes. Notes: Nominal Commodity Price Index and Nominal Commodity Price indices for major subgroups, metals, food, energy and materials. The quarterly series extracted from IFS database.

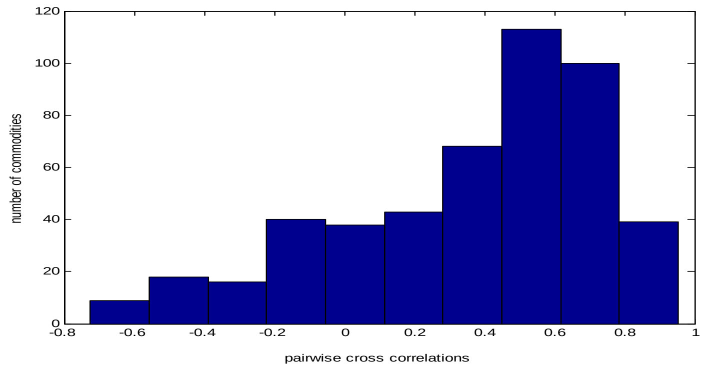

Figure 10: Cross Correlations. Notes: Histogram lists all the pairwise cross correlations across 42 non-energy nominal commodity prices.

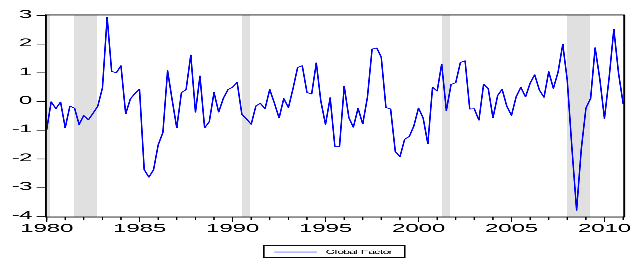

Figure 11: Estimated Median Global Factor. Notes: Shaded areas represent the NBER recessions.

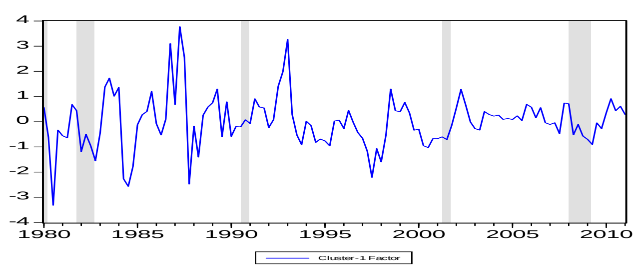

Figure 12: Estimated Median Cluster-1 ("Timber"). Notes: Shaded areas represent the NBER recessions. This cluster is dominantly made of logs and wood, lamb meat weakly belong to this cluster.

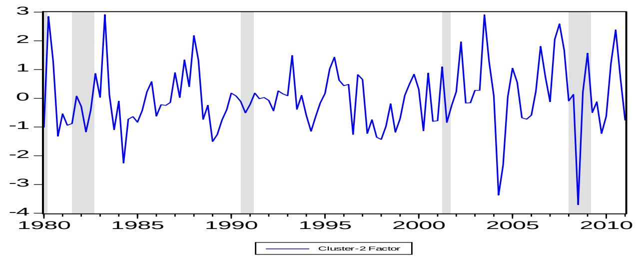

Figure 13: Estimated Median Cluster-2 ("Grains, Oil"). Notes: Shaded areas represent the NBER recessions. This cluster is dominantly made of grains (except rice) and vegetable oils (except olive oil). Some other food products and cotton weakly belong to this cluster.

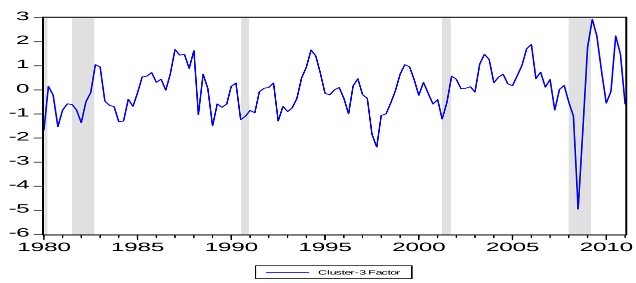

Figure 14: Estimated Median Cluster-3 ("Mixed"). Notes: Shaded areas represent the NBER recessions. This cluster consists of metals (except iron), agricultural raw materials (except cotton and timber) and some food products.

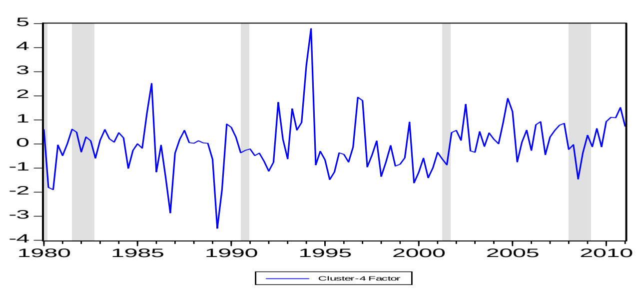

Figure 15: Estimated Median Cluster-4 ("Coffee"). Notes: Shaded areas represent the NBER recessions. This cluster is dominantly made of coffee. Iron, shrimp and swine also weakly belong to this cluster with the degrees in the written order.

APPENDIX A

Appendix For Chapter 2

A.1 Technical Details

The following subsections contain the draws omitted from the main text and more details on the draw of the factors.

A.1.1 Generating $\boldsymbol{\sigma}^2|\boldsymbol{\Theta}_{-\boldsymbol{\sigma}^2}, \mathbf{F}, \mathbf{Y}$

Next, we draw σ_n^2 conditional on $\mathbf{Y}$, $\mathbf{F}$, and $\Theta_{-\sigma_n^2}$. Let $\overline{\mathbf{Y}}_n$ reflect the quasi-difference of the vector $\mathbf{Y}_n$ and $\overline{\mathbf{X}}_n$ reflect the quasi-difference of the vector $\mathbf{X}_n = [\mathbf{1}_T, \mathbf{f}_0, \widetilde{\mathbf{F}}\gamma_n]$, conditioned on the accepted value of γ_n. The innovation variance is a straightforward draw from an inverse gamma posterior

$$\sigma_n^{-2}|\mathbf{Y},\mathbf{X},\Theta_{-\sigma_n^2} \sim \boldsymbol{\Gamma}\left(\frac{\nu_0+T}{2}, \frac{\Upsilon_0 + \left(\overline{\mathbf{Y}}_n - \overline{\mathbf{X}}_n\beta_n\right)'\left(\overline{\mathbf{Y}}_n - \overline{\mathbf{X}}_n\beta_n\right)}{2}\right).$$

A.1.2 Generating $\boldsymbol{\psi}|\boldsymbol{\Theta}_{-\boldsymbol{\psi}}, \mathbf{F}, \mathbf{Y}$

The set of serial correlation coefficients, $\boldsymbol{\psi}_n = \left[\boldsymbol{\psi}_{n1}, .., \boldsymbol{\psi}_{np_\varepsilon}\right]$, can be sampled, conditional on $\boldsymbol{\Theta}_{-\boldsymbol{\psi}}$, $\mathbf{F}$,and $\mathbf{Y}$, in an MH step [Chib (1993); Chib and Greenberg (1994)]. We draw a candidate $\boldsymbol{\psi}_n^*$ from the proposal density: $\boldsymbol{\psi}_n^* \sim N\left(\mathbf{w}_n, \mathbf{W}_n\right)$, where

$$\mathbf{W}_n = \mathbf{W}_0 + \sigma_n^{-2}\boldsymbol{\varepsilon}_n'\boldsymbol{\varepsilon}_n,$$

$$\mathbf{w}_n = \mathbf{W}_n\left(\mathbf{W}_0^{-1}\mathbf{w}_0 + \sigma_n^{-2}\boldsymbol{\varepsilon}_n'\widehat{\varepsilon}_{n0}\right),$$

$\boldsymbol{\varepsilon}_n = [\widehat{\varepsilon}_{n1}, ..., \widehat{\varepsilon}_{np_\varepsilon}]$, $\widehat{\varepsilon}_{nk} = [\varepsilon_{n,p_\varepsilon+1-k}, ..., \varepsilon_{n,T-k}]'$, and

$$\varepsilon_{nt} = y_{nt} - \beta_{n0} - \beta_{nG} f_t - \sum_i^M \gamma_{ni}\beta_{ni}F_{it}.$$

Once we have obtained a candidate, it is accepted with probability

$$A_{n,\psi} = \min\left\{\frac{\Psi(\boldsymbol{\psi}_n^*)}{\Psi(\boldsymbol{\psi}_n)}, 1\right\}, \tag{15}$$

where

$$\Psi(\boldsymbol{\psi}_n) = |\Sigma_n|^{-1/2}\exp\left[-\frac{1}{2\sigma_n^2}(\mathbf{Y}_n - \mathbf{X}_n\beta_n)'\Sigma_n^{-1}(\mathbf{Y}_n - \mathbf{X}_n\beta_n)\right],$$

the quantity Σ_n is the function defined above of either the candidate or last iteration value of $\boldsymbol{\psi}_n$, and the denominator in (15) is computed with the value from the previous iteration. Thus, the acceptance probability is the ratio of the pseudolikelihoods for the candidate and the past accepted draw.

A.1.3 Generating $\boldsymbol{\phi}|\boldsymbol{\Theta}_{-\phi}, \mathbf{F}, \mathbf{Y}$

The draw for $\boldsymbol{\phi}_i = \left[\boldsymbol{\phi}_{i0}, \boldsymbol{\phi}_{i1}, .., \boldsymbol{\phi}_{ip_F}\right]$ for $i = 1, ..., M$ is similar to the draw of $\boldsymbol{\psi}_n$ the subsection. The candidate factor AR coefficients are drawn from the proposal: $\boldsymbol{\phi}_i^* \sim N(\mathbf{v}_i, \mathbf{V}_i)$, where

$$\mathbf{v}_i = \mathbf{V}_i\left(\mathbf{V}_0^{-1}\mathbf{v}_0 + \mathbf{e}_i'\widehat{e}_{i0}\right),$$

$$\mathbf{V}_i = \mathbf{V}_0 + \boldsymbol{\varepsilon}_i'\boldsymbol{\varepsilon}_i,$$

$\mathbf{e}_i = [\widehat{e}_{i1}, ..., \widehat{e}_{ip_F}]$, $\widehat{e}_{ik} = [e_{i,p_F+1-k}, ..., e_{i,T-k}]'$, and

$$e_{it} = F_{it} - \phi_i (L) F_{it-1}.$$

As before, the candidate is accepted with a probability

$$A_{n,\phi} = \min \left\{ \frac{\Psi (\boldsymbol{\phi}_i^*)}{\Psi \left(\boldsymbol{\phi}_i^{[g-1]} \right)}, 1 \right\},$$

where the pseudolikelihood, $\Psi (\boldsymbol{\phi}_i)$, is

$$\Psi (\boldsymbol{\phi}_i) = |\Omega_i|^{-1/2} \exp \left[-\frac{1}{2} \widehat{e}_{i0}' \Omega_i^{-1} \widehat{e}_{i0} \right]$$

with Ω_i defined as above. The same algorithm is repeated to obtain the AR parameters for the global factor.

A.1.4 Generating $\mathbf{F}|\boldsymbol{\Theta}, \boldsymbol{\phi}, \mathbf{Y}$

The factors are generated recursively from posterior distributions obtained from the modified Kalman filter described in Kim and Nelson (1999). For notational simplicity, we describe the case in which the factors and observation errors have the same lag order, p. Let $\widetilde{y}_t = y_t - \beta_0$ represent the conditionally demeaned data and $\rho_t = [F_t', \varepsilon_t', ..., F_{t+1-p}', \varepsilon_{t+1-p}']'$ be the state vector. Define $\overline{N} = M + N + 1$. Then, the state-space representation is described by

$$\widetilde{y}_t = H \rho_t, \tag{16}$$

where

$$H = \left[\begin{array}{cccc} \boldsymbol{\beta}_G & \gamma\beta & \mathbf{1}_{N \times N} & \mathbf{0}_{N \times \overline{N}(p-1)} \end{array} \right];$$

$$\rho_t = G\rho_{t-1} + \hat{u}_t; \tag{17}$$

$$G = \begin{bmatrix} G_1 & \cdots & G_p \\ \mathbf{1}_{\overline{N}(p-1)\times\overline{N}(p-1)} & & \mathbf{0}_{\overline{N}(p-1)\times\overline{N}} \end{bmatrix};$$

$G_i = diag\left[\phi_{G,i}, \phi_{1,i}, ..., \phi_{M,i}, \psi_{1,i}, ..., \psi_{N,i}\right]$; $\hat{u}_t = \left[e_t', \epsilon_t', \mathbf{0}'_{\overline{N}\times(p-1)}\right]'$; $E_t\left[\hat{u}_t\hat{u}_t'\right] = Q$; and

$$Q = \begin{bmatrix} diag\left[\omega_G^2, \omega_1^2, ..., \omega_M^2, \sigma_1^2, ..., \sigma_N^2\right] & \mathbf{0}_{\overline{N}\times\overline{N}(p-1)} \\ \mathbf{0}_{\overline{N}(p-1)\times\overline{N}} & \mathbf{0}_{\overline{N}(p-1)\times\overline{N}(p-1)} \end{bmatrix}.$$

The Kalman filter sequentially updates the linear projection for the system (16) and (17) above. Given an initialization for the state vector, $\rho_{0|0}$, and for the unconditional density of the state vector; $P_{0|0}$, the Kalman filter computes the prediction density by propagating the state vector through the transition equation, (17), to obtain an estimate of the state vector at period t conditional on information available at time $t-1$:

$$\begin{aligned} \rho_{t|t-1} &= G\rho_{t-1|t-1}, \\ P_{t|t-1} &= GP_{t-1|t-1}G' + Q. \end{aligned}$$

The filter then updates this prediction given the new (observable) information realized at time t as:

$$\begin{aligned} \rho_{t|t} &= \rho_{t|t-1} + P_{t|t-1}H'(H'P_{t|t-1}H + R)^{-1}(\widetilde{y}_t - H\rho_{t|t-1}), \\ P_{t|t} &= P_{t|t-1} - P_{t|t-1}H'(H'P_{t|t-1}H + R)^{-1}HP_{t|t-1}. \end{aligned}$$

The factors are then sampled recursively from a Gaussian distribution with one additional step. Because Q is singular and we are only interested in drawing the factors, Kim and Nelson (1999) show that the dimensions of Q, G, and the state variable must be reduced. They define Q^* as the upper $(M+1) \times (M+1)$ submatrix of Q and

$$G^* = \begin{bmatrix} G_1 & \mathbf{0}_{(M+1 \times N)} & G_2 & \mathbf{0}_{(M+1 \times N)} & \cdots & G_p & \mathbf{0}_{(M+1 \times N)} \end{bmatrix},$$

where G_i is defined above.

The last iteration of the Kalman filter yields $\rho_{T|T}$ and $P_{T|T}$ from which we can draw F_T. Then, we recursively draw F_{t-1} from $N\left(\rho_{t|t,\rho^*_{t+1}}, P_{t|t,\rho^*_{t+1}}\right)$, where

$$\begin{aligned} \rho_{t|t,\rho^*_{t+1}} &= \rho_{t|t} + P'_{t|t}G^{*\prime}(G^*P_{t|t}G^{*\prime} + Q^*)^{-1}(\rho^*_{t+1} - G^*\rho_{t|t}), \\ P_{t|t,\rho^*_{t+1}} &= P_{t|t} - P_{t|t}G^{*\prime}(G^*P_{t|t}G^{*\prime} + Q^*)^{-1}G^*P_{t|t}. \end{aligned}$$

A.1.5 Computing the Bayes Factors

Chib (1995) uses the basic marginal likelihood identity to approximate the marginal likelihood using the output from the Gibbs sampler:

$$\ln \widehat{m}(\mathbf{Y}) = \ln f(\mathbf{Y}|\mathbf{\Theta}^*) + \ln \pi(\mathbf{\Theta}^*) - \ln \widehat{\pi}(\mathbf{\Theta}^*|\mathbf{Y}), \tag{18}$$

where Θ is the vector of model parameters, $\ln \widehat{m}(\mathbf{Y})$ is the log marginal likelihood, $\ln f(\mathbf{Y}|\mathbf{\Theta}^*)$ is the log likelihood evaluated at a given $\mathbf{\Theta} = \mathbf{\Theta}^*$, $\ln \pi(\mathbf{\Theta}^*)$ is the log of the prior evaluated at $\mathbf{\Theta}^*$, $\ln \widehat{\pi}(\mathbf{\Theta}^*|\mathbf{Y})$ is the posterior ordinate, and $\mathbf{\Theta}^*$ is any high density value of $\mathbf{\Theta}$ (e.g., a modal point).

The posterior ordinate $\ln \widehat{\pi}(\mathbf{\Theta}^*|\mathbf{Y})$ can be computed by expanding the expression

$\widehat{\pi}(\mathbf{\Theta}^*|\mathbf{Y})$ as

$$\widehat{\pi}(\mathbf{\Theta}^*|\mathbf{Y}) = \pi(\Theta_1^*|\mathbf{Y}) \times \pi(\Theta_2^*|\mathbf{Y}, \Theta_1^*) \times ... \times \pi\left(\Theta_N^*|\mathbf{Y}, \Theta_1^*, ..., \Theta_{N-1}^*\right),$$

where here N represents the number of blocks in the sampler. The first term can be obtained from the initial run of the Gibbs sampler. A typical term for block n in the above representation is $\pi\left(\Theta_n^*|\mathbf{Y}, \Theta_1^*, ..., \Theta_{n-1}^*\right)$, which can be estimated by additional sampling of $\{\Theta_{n+1}..., \Theta_N, \mathbf{F}\}$ holding constant $\{\Theta_1^*, ..., \Theta_n^*\}$. In general, the estimate of the posterior ordinate, $\widehat{\pi}(\Theta_n^*|\mathbf{Y})$, is then

$$\widehat{\pi}(\Theta_n^*|\mathbf{Y}) = \frac{1}{G}\sum_{g=1}^{G} \pi\left(\Theta_n^*|\mathbf{Y}, \Theta_1^*, ..., \Theta_{n-1}^*, \Theta_{n+1}^{(g)}, ..., \Theta_N^{(g)}, F_0^{(g)}, F_1^{(g)}, ..., F_M^{(g)}, f_1^{(g)}, ..., f_N^{(g)}\right).$$

For the blocks requiring an MH step, we employ the method of Chib and Jeliazkov (2001), which computes the posterior ordinate as proportional to the sum of the candidates weighted by their appropriate acceptance probabilities.

A.2 Data

The indices in (1), (2) and (3) above are obtained from Baxter and Kouparitsas (2003). Openness is measured in constant prices (percent in 2005 constant prices); exports + imports divided by RGDPL and represents the total trade as a percentage of GDP. Investment share of real GDP per capita (RGDPL) is defined in 2005 constant prices. Real GDP chain per equivalent adult in 2005 constant prices computes output with weights of 1.0 to all persons over 15 and 0.5 for those under age 15.

Country Listing

Oceania	Latin America		Europe		Africa	Asia 1	Asia 2
Australia	C. Rica	Bolivia	France	Italy	Cameroon	Bangladesh	H.K.
N.Z.	D. Repub.	Brazil	Austria	Lux.	I. Coast	India	Japan
	El Salv.	Chile	Belgium	Neth.	Kenya	Indonesia	S. Korea
N. America	Guat.	Columbia	Denmark	Norway	Morocco	Pakistan	Malaysia
Canada	Hond.	Ecuador	Finland	Portugal	Senegal	Philippines	Sing.
U.S.	Jamaica	Paraguay	Germany	Spain	S. Africa	Sri Lanka	Thailand
Mexico	Panama	Peru	Greece	Sweden	Zimbabwe		
	Trinidad	Uruguay	Iceland	Switz.			
	Argentina	Venezuela	Ireland	U.K.			

Table 15: Country Listing. Notes: Regions in bold reflect the groupings imposed in Kose, Otrok, and Whiteman (2003).

A.3 Full Posterior Distributions

Posterior Means and Coincidence Intervals for Cluster Loadings - I

	β_1			β_2			β_3		
	Mean	16th-84th		Mean	16th-84th		Mean	16th-84th	
Argentina	-0.56	-0.89	-0.27	0.57	0.57	0.57	1.19	0.79	1.59
Australia	NA	NA	NA	1.24	1	1.5	NA	NA	NA
Austria	-1.06	-1.25	-0.87	1.42	1.31	1.64	NA	NA	NA
Bangladesh	0.14	-0.27	0.54	0.27	-0.18	0.68	-0.06	-0.43	0.33
Belgium	-1.3	-1.49	-1.13	1.18	0.69	1.65	0.93	0.64	1.31
Bolivia	-0.14	-0.43	0.23	0.21	-0.18	0.59	0.74	0.36	1.09
Brazil	-1.23	-1.63	-0.85	-0.06	-0.57	0.47	1.23	0.82	1.63
Cameroon	-0.01	-0.55	0.5	0.68	0.15	1.2	0.39	-0.14	0.89
Canada	-1.1	-1.37	-0.74	1.69	1.46	1.93	0.99	0.83	1.27
Chile	-0.08	-0.53	0.48	0.15	-0.46	0.75	0.88	0.44	1.34
Colombia	-0.34	-0.57	-0.2	0.3	0.3	0.3	0.64	0.45	0.82
Costa Rica	-0.65	-0.75	-0.52	0.98	0.61	1.22	0.93	0.66	1.19
Ivory Coast	-0.87	-1.36	-0.35	0.54	0.01	1.08	-0.06	-0.56	0.44
Denmark	-1.02	-1.31	-0.73	1.15	0.87	1.43	NA	NA	NA
Dom. Republic	-0.56	-1.02	-0.14	0.93	0.47	1.39	0.85	0.44	1.26
Ecuador	-0.34	-0.76	0.07	-0.31	-0.6	0.05	1.01	0.63	1.38
El Salvador	-0.32	-0.6	-0.03	0.78	0.4	1.2	0.74	0.4	1.1
Finland	-1.78	-2.09	-1.47	NA	NA	NA	NA	NA	NA
France	-1.26	-1.47	-1.06	1.33	0.68	1.94	0.77	0.05	1.49
Germany	-0.55	-0.76	-0.33	0.36	0.15	0.59	0.2	-0.01	0.38
Greece	-1.03	-1.47	-0.61	0.87	0.51	1.27	0.36	-0.03	0.74
Guatemala	-0.47	-0.7	-0.17	0.1	-0.26	0.45	0.65	0.37	0.92
Honduras	0.28	-0.11	0.54	0.2	-0.27	0.69	0.88	0.57	1.19
Hong Kong	-0.65	-1.15	-0.16	0.96	0.43	1.51	0.92	0.44	1.43
Iceland	-1.07	-1.56	-0.58	1.25	0.74	1.76	1	0.53	1.49
India	-0.22	-0.61	0.14	0.56	0.14	0.97	0.11	-0.29	0.49
Indonesia	0.03	-0.36	0.43	0.16	-0.33	0.64	0.14	-0.24	0.52
Ireland	-0.92	-1.31	-0.52	0.8	0.37	1.23	0.3	-0.17	0.77
Italy	-1.34	-1.56	-1.11	NA	NA	NA	NA	NA	NA
Jamaica	0.05	-0.39	0.49	-0.01	-0.47	0.45	-0.19	-0.64	0.28

Table 16: Posterior Means and Coincidence Intervals for Cluster Loadings - I. Notes: The columns show posterior means, 16th and 84th percentiles for the cluster loadings.

Posterior Means and Coincidence Intervals for Cluster Loadings - II

	β_1			β_2			β_1		
	Mean	16th-84th		Mean	16th-84th		Mean	16th-84th	
Japan	-0.74	-1.05	-0.42	0.03	-0.27	0.37	-0.04	-0.48	0.45
Kenya	-0.34	-0.7	0.04	0.54	0.15	0.94	0.34	0.03	0.68
Korea	-0.27	-0.69	0.13	0.73	0.27	1.19	-0.51	-0.87	-0.17
Luxemburg	-1.41	-1.82	-1.02	0.94	0.55	1.43	0.11	-0.34	0.56
Malaysia	-0.12	-0.5	0.27	0.24	-0.16	0.65	0.43	-0.04	0.91
Mexico	-0.53	-0.93	-0.06	0.56	0.01	1.1	0.87	0.5	1.26
Morocco	-0.53	-1.11	0.06	-0.4	-0.92	0.13	-0.22	-0.77	0.31
Netherlands	-0.95	-1.17	-0.74	0.9	0.65	1.15	0.66	0.66	0.66
New Zealand	-0.42	-0.9	-0.02	1.13	0.74	1.52	-0.16	-0.16	-0.16
Norway	-0.65	-0.87	-0.41	0.32	0.02	0.67	0.75	0.44	1.1
Pakistan	-0.51	-0.77	-0.25	0.26	-0.05	0.56	0.33	0.03	0.62
Panama	-0.45	-0.97	-0.07	-0.4	-0.45	-0.35	0.7	0.32	1.06
Paraguay	-0.32	-0.73	-0.01	-0.08	-0.26	0.06	0.52	0.14	0.92
Peru	-0.74	-1.22	-0.24	0.77	-0.09	1.56	1.1	0.65	1.55
Philippines	-0.4	-0.77	0	0.15	-0.38	0.67	0.76	0.44	1.09
Portugal	-1.55	-1.92	-1.19	NA	NA	NA	-0.03	-0.27	0.15
Senegal	0.19	-0.14	0.54	-0.15	-0.53	0.24	0.07	-0.21	0.34
Singapore	-0.19	-0.67	0.3	0.2	-0.39	0.8	0.41	-0.11	0.89
South Africa	-0.16	-0.5	0.16	0.05	-0.24	0.35	0.08	-0.37	0.43
Spain	-1.17	-1.45	-0.9	NA	NA	NA	-0.37	-0.73	-0.09
Sri Lanka	-0.89	-1.21	-0.57	0.14	-0.24	0.49	0.39	0.08	0.75
Sweden	-1.15	-1.35	-0.95	NA	NA	NA	NA	NA	NA
Switzerland	-1.17	-1.37	-0.97	NA	NA	NA	NA	NA	NA
Thailand	0.46	0.09	0.82	-0.46	-0.87	-0.05	-0.15	-0.53	0.18
Trinidad&Tobago	-0.98	-1.54	-0.46	-0.07	-0.68	0.55	1.02	0.42	1.53
United Kingdom	NA	NA	NA	1.19	0.96	1.43	NA	NA	NA
United States	-1.4	-1.99	-1.03	1.72	1.45	1.98	NA	NA	NA
Uruguay	0.07	-0.32	0.48	0.17	-0.4	0.7	1.12	0.6	1.61
Venezuela	-0.25	-0.95	0.04	NA	NA	NA	1.36	0.91	1.8
Zimbabwe	-0.58	-1.26	0.1	-0.52	-1.29	0.3	-0.2	-0.97	0.65

Table 17: Posterior Means and Coincidence Intervals for Cluster Loadings - II. Notes: The columns show posterior means, 16th and 84th percentiles for the cluster loadings. 'NA' represents that the observation never chose that cluster and its corresponding loading was not sampled.

Posterior Means and Coincidence Intervals for Parameters - I

	β_0			β_G			σ^2		
	Mean	16th-84th		Mean	16th-84th		Mean	16th-84th	
Argentina	0.05	-0.71	0.82	0.07	-0.4	0.54	13.13	10.17	16.07
Australia	0.6	0.04	1.15	-0.33	-0.6	-0.06	2.79	2.17	3.4
Austria	0.05	-0.37	0.49	-0.62	-0.85	-0.39	1.7	1.31	2.09
Bangladesh	0.78	0	1.54	0.14	-0.29	0.58	13.73	10.88	16.52
Belgium	-0.41	-0.81	-0.02	-0.66	-0.87	-0.45	0.79	0.49	0.91
Bolivia	-0.31	-0.98	0.36	-0.27	-0.63	0.08	6.13	4.79	7.47
Brazil	-0.02	-0.83	0.8	-0.49	-0.95	-0.02	9.82	7.69	11.92
Cameroon	0.43	-0.44	1.33	0.71	0.15	1.27	20.92	16.61	25.3
Canada	0.23	-0.34	0.81	-0.36	-0.63	-0.09	0.72	0.38	0.97
Chile	0.04	-0.82	0.89	-1.07	-1.56	-0.59	12.43	9.7	15.14
Colombia	0.38	-0.1	0.86	-0.53	-0.75	-0.3	1.6	1.2	1.99
Costa Rica	0.72	0.11	1.34	-0.15	-0.51	0.21	5.52	4.26	6.76
Ivory Coast	-0.25	-1.1	0.59	-0.21	-0.71	0.29	24.75	19.7	29.8
Denmark	0.37	-0.25	1	-0.46	-0.78	-0.14	4.38	3.45	5.31
Dom. Republic	0.99	0.22	1.79	-0.53	-1.02	-0.05	16.69	15.64	23.76
Ecuador	0.04	-0.74	0.8	-0.19	-0.61	0.23	8.78	6.87	10.64
El Salvador	0.17	-0.48	0.81	-0.11	-0.44	0.23	5.23	4.06	6.37
Finland	1.16	0.47	1.86	0.58	0.2	0.95	3.93	2.86	4.98
France	-0.25	-0.68	0.17	-0.46	-0.68	-0.24	0.94	0.66	1.16
Germany	-0.28	-0.9	0.33	-0.67	-0.89	-0.45	1.4	1.08	1.72
Greece	0.23	-0.59	1.08	-0.79	-1.24	-0.35	9.42	7.45	11.4
Guatemala	-0.39	-1.08	0.3	-0.78	-1.06	-0.5	3.13	2.39	3.86
Honduras	-0.09	-0.76	0.6	-0.12	-0.51	0.27	10.27	8.01	12.52
Hong Kong	0.64	-0.21	1.51	-1.61	-2.1	-1.12	11.66	9.13	14.21
Iceland	0.87	0.02	1.7	0	-0.55	0.54	22.09	17.54	26.7
India	1.60	0.79	2.4	-0.19	-0.6	0.21	8.56	6.79	10.28
Indonesia	0.23	-0.53	1	-1.71	-2.1	-1.32	6.66	5.07	8.26
Ireland	0.50	-0.35	1.36	-0.99	-1.41	-0.57	7.3	5.78	8.82
Italy	-0.60	-1.1	-0.1	-0.77	-1.04	-0.5	1.87	1.42	2.32
Jamaica	0.10	-0.73	0.94	-0.31	-0.77	0.16	12.64	10.08	15.22

Table 18: Posterior Means and Coincidence Intervals for Parameters - I. Notes: The columns show posterior means, 16th and 84th percentiles for the parameters: intercept, global factor loading and the idiosyncratic variance.

Posterior Means and Coincidence Intervals for Parameters - II

	β_0			β_G			σ^2		
	Mean	16th-84th		Mean	16th-84th		Mean	16th-84th	
Japan	0.22	-0.68	1.13	-1.03	-1.4	-0.65	4.72	3.68	5.77
Kenya	-0.11	-0.87	0.63	-0.01	-0.38	0.36	12.87	10.22	15.49
Korea	1.43	0.56	2.28	-1.95	-2.38	-1.52	8.9	6.73	11.06
Luxemburg	0.1	-0.74	0.91	-0.58	-1.02	-0.14	7.63	6.03	9.21
Malaysia	0.73	-0.01	1.46	-1.81	-2.19	-1.43	4.75	3.49	6.05
Mexico	0.24	-0.46	0.94	-0.49	-0.91	-0.07	9.85	7.77	11.9
Morocco	0.21	-0.78	1.2	0	-0.59	0.6	23.83	19.09	28.6
Netherlands	0.03	-0.47	0.52	-0.44	-0.7	-0.19	2.22	1.73	2.7
New Zealand	0.19	-0.56	0.93	-0.16	-0.56	0.25	7.94	6.29	9.54
Norway	1.52	0.97	2.08	-0.24	-0.51	0.02	2.93	2.3	3.55
Pakistan	1.56	0.85	2.29	-0.06	-0.39	0.26	4.77	3.79	5.73
Panama	1.13	0.42	1.84	-0.41	-0.84	0.01	15.03	11.89	18.15
Paraguay	0.34	-0.45	1.13	0.01	-0.44	0.47	9.83	7.82	11.79
Peru	0.49	-0.3	1.3	0.2	-0.32	0.71	13.54	10.5	16.62
Philippines	-0.02	-0.73	0.7	-0.62	-1.09	-0.15	8.99	7.09	10.88
Portugal	-0.64	-1.4	0.11	-1.14	-1.58	-0.69	6.84	5.29	8.35
Senegal	0.82	0.12	1.53	0.81	0.44	1.2	13.2	10.33	15.96
Singapore	0.60	-0.23	1.42	-2.17	-2.7	-1.64	13.83	10.29	17.31
South Africa	0.85	0.1	1.6	-0.12	-0.45	0.22	3.46	2.76	4.17
Spain	0.42	-0.54	1.37	-0.18	-0.5	0.13	3.14	2.45	3.81
Sri Lanka	0.74	-0.05	1.53	-0.44	-0.8	-0.07	6.4	5.06	7.77
Sweden	1.05	0.5	1.58	0.39	0.09	0.68	1.55	1.12	1.97
Switzerland	-0.52	-0.99	-0.05	-0.27	-0.51	-0.03	1.47	1.14	1.8
Thailand	1.70	0.93	2.49	-1.58	-2.01	-1.16	6.28	4.59	7.94
Trinidad&Tobago	0.1	-0.84	1.03	-1.16	-1.78	-0.53	28.9	22.77	35.02
United Kingdom	0.29	-0.25	0.84	-0.46	-0.71	-0.22	1.71	1.32	2.1
United States	-0.66	-1.25	-0.05	-0.8	-1.07	-0.5	1.07	0.71	1.41
Uruguay	-0.13	-0.92	0.68	-0.36	-0.87	0.15	13.85	10.89	16.83
Venezuela	-0.79	-1.6	0.03	-0.18	-0.69	0.32	21.42	16.7	26.08
Zimbabwe	-0.12	-1.05	0.81	0.05	-0.69	0.78	86.19	69.1	103.16

Table 19: Posterior Means and Coincidence Intervals for Parameters - II. Notes: The columns show posterior means, 16th and 84th percentiles for the parameters: intercept, global factor loading and the idiosyncratic variance.

	Cluster 1	Cluster 2	Cluster 3		Cluster 1	Cluster 2	Cluster 3
Argentina	0	0	1	Japan	0.95	0.04	0.01
Australia	0	1	0	Kenya	0.07	0.89	0.04
Austria	1	0	0	Korea	0.43	0.38	0.19
Bangladesh	0.27	0.25	0.48	Luxemburg	0.95	0.03	0.02
Belgium	0.98	0.02	0	Malaysia	0.09	0.9	0.01
Bolivia	0.01	0	0.99	Mexico	0.03	0	0.96
Brazil	0.69	0.02	0.28	Morocco	0.15	0.04	0.81
Cameroon	0.07	0.79	0.14	Netherlands	1	0	0
Canada	0.01	0.99	0	New Zealand	0.01	0.99	0
Chile	0.02	0	0.98	Norway	0.9	0.03	0.07
Colombia	0	0	1	Pakistan	0.52	0.35	0.13
Costa Rica	0	0	1	Panama	0.01	0	0.99
Ivory Coast	0.23	0.66	0.1	Paraguay	0.01	0	0.99
Denmark	0.2	0.8	0	Peru	0.03	0	0.97
Dom. Rep.	0.07	0.02	0.91	Philippines	0.06	0.05	0.88
Ecuador	0.02	0	0.98	Portugal	1	0	0
El Salvador	0.03	0.03	0.94	Senegal	0.09	0.29	0.62
Finland	1	0	0	Singapore	0.45	0.42	0.13
France	0.99	0.01	0	South Africa	0.03	0.97	0
Germany	0.91	0.06	0.04	Spain	1	0	0
Greece	0.88	0.07	0.05	Sri Lanka	0.9	0.02	0.08
Guatemala	0.02	0	1	Sweden	1	0	0
Honduras	0	0	0	Switzerland	1	0	0
Hong Kong	0.05	0.95	0	Thailand	0.62	0.37	0.02
Iceland	0.59	0.1	0.31	Trinidad & Tobago	0.55	0.44	0.1
India	0.12	0.85	0.03	United Kingdom	0	1	0
Indonesia	0.21	0.62	0.17	United States	0	1	0
Ireland	0.63	0.35	0.01	Uruguay	0.02	0.01	0.97
Italy	1	0	0	Venezuela	0	0	1
Jamaica	0.47	0.48	0.04	Zimbabwe	0.19	0.77	0.04

Table 20: Cluster Probabilities - Posterior Means. Notes: Each value represents the posterior means of cluster membership indicator for each country.

APPENDIX B

Appendix For Chapter 3

B.1 Estimation Details

Following FOS (2012) and Kose et al. (2003), I estimate the model presented in equation (9) via Gibbs sampling. In particular, I utilize Metropolis-Hastings in Gibbs sampling to draw from the joint posterior distribution of the factors and the model's parameters. Given an initial draw of model's parameters, the factors can be extracted using Kalman filters based on Kim and Nelson (1999). In the next step, taking these factors as given, one can sample from the conditional density of the parameters. Once the parameters are known, Kalman filtering technique is applied again to extract the factors. Sampling from the conditional densities of the parameters and the factors is repeated many times. This is known as Gibbs sampling and under the regulatory assumptions (see Chip and Greenberg, 1995) these sequence of draws from the conditional distributions converge to the joint posterior density of the entire system, independent of the initial values selected.

To describe the sampler fully, let $\mathbf{Y}$ represent the data, $\mathbf{\Theta}$ represent the full set of model parameters and let $\mathbf{F}$ represent the factors. We can define the set of blocks of parameters to be estimated in the sampler as: (1) the set of intercepts, $\boldsymbol{\alpha}_i$ and global factor loadings, β_{i0} collected in the set $\boldsymbol{\rho}=\{\alpha_i, \beta_{i0}\}$; (2) the set of innovation variances, $\boldsymbol{\sigma}^2 = \{\boldsymbol{\sigma}_i^2\}$; (3) the set of autoregressive parameters for the factors, $\phi = \{\phi_0, ., \phi_J\}$, (4) the sectoral factor loadings $\boldsymbol{\beta} = \{\beta_{i,j}\}$ joint with the group probabilities $\boldsymbol{\gamma} = \{\boldsymbol{\gamma}_{i,j}\}$, (5), the set of factors, $\mathbf{F} = \{\mathbf{F}_0, \mathbf{F}_j\}$ and (6) the set of autoregressive parameters for the factors, $\varphi = \{\varphi_1, ., \varphi_I\}$

The steps of the Gibbs algorithm to sample from the joint distribution of $\mathbf{\Theta}, \mathbf{F}$ are given as follows:

Step 1: Specify starting values $\mathbf{\Theta}^{(0)},\ \mathbf{F}^{(0)}$ and set $n = 0$.

Step 2: Simulate the unknown variables;

2.1: Sample $\boldsymbol{\rho}^{(n+1)}$ from $p(\ \rho|(\sigma^2)^{(n)},\ \phi^{(n)},\ (\beta,\gamma)^{(n)},\ \varphi^{(n)}, F^{(n)},\ Y)$,

2.2: Sample $(\boldsymbol{\sigma}^2)^{(n+1)}$ from $p(\ \sigma^2|\rho^{(n+1)},\ \phi^{(n)},\ (\beta,\gamma)^{(n)}, \varphi^{(n)},\ F^{(n)},\ Y)$,

2.3: Sample $(\boldsymbol{\beta},\boldsymbol{\gamma})^{(n+1)}$ from $p(\ \boldsymbol{\beta},\boldsymbol{\gamma}|\rho^{(n+1)},\ (\sigma^2)^{(n+1)},\ \phi^{(n)},\ \varphi^{(n)}, F^{(n)}, Y)$,

2.4: Sample $\phi^{(n+1)}$ from $p(\ \boldsymbol{\gamma}|\rho^{(n+1)},\ (\sigma^2)^{(n+1)},\ (\beta,\gamma)^{(n+1)},\ \varphi^{(n)}, F^{(n)}, Y)$,

2.5 Sample $\varphi^{(n+1)}$ from $p(\ \boldsymbol{\varphi}|\rho^{(n+1)},\ (\sigma^2)^{(n+1)},\ (\beta,\gamma)^{(n+1)}, \phi^{(n+1)},\ F^{(n)}, Y)$,

2.6 Apply Kalman Filter and sample $\mathbf{F}^{(n+1)}$.

Step 3: Set $n = n + 1$ and go to step 2.

This iteration loop is repeated 30,000 times and the initial 25,000 draws are discarded to allow for convergence. To initialize the sampler, I generate factors from a uniform normal distribution, and run the regressions of (9) and (10) separately. The coefficient estimates of factor loadings, factor AR parameters and variances for measurement errors are then used to start the sampler.

B.1.1 The Prior Distributions

The prior distributions and their corresponding hyperparameters are given below:

Priors for Estimation - II

Parameter	Prior Distribution	Hyperparameters	
$\rho_i = \left[\alpha_i, \beta_{i,0}\right]'$	$N(\mathbf{r},\mathbf{R})$	$\mathbf{r} = \mathbf{0}_2\ ;\ \mathbf{R} = \mathbf{I}_2$	$\forall\ i$
$\beta_{i,j}$	$N(\mathbf{b},\mathbf{B})$	$\mathbf{b} = \mathbf{0}\ ;\ \mathbf{B} = \mathbf{1}$	$\forall\ i,\ j$
σ_i^2	$I\Gamma\left(\frac{\nu}{2},\frac{\delta}{2}\right)$	$\nu = 6\ ;\ \delta = 0.1$	$\forall\ i$
$\gamma_{i,j}$	$\mathbf{Uniform}(\kappa)$	$\kappa_{ij} = \frac{1}{J}$	$\forall\ i, j$
$\boldsymbol{\phi}$	$N(\boldsymbol{\eta},\boldsymbol{\Phi})$	$\boldsymbol{\eta} = \mathbf{0}_{q^f},\ \boldsymbol{\Phi} = \frac{1}{2}\mathbf{I}_{q^f}$	$\forall\ j$
φ	$N\left(\bar{\boldsymbol{\mu}},\bar{\boldsymbol{\Upsilon}}\right)$	$\bar{\boldsymbol{\mu}} = \mathbf{0}_{q^\varepsilon},\ \bar{\boldsymbol{\Upsilon}} = \frac{1}{2}\mathbf{I}_{q^\varepsilon}$	$\forall\ i$

Table 21: Priors for Estimation - II.

Note that the cluster membership indicator has a uniform prior over all clusters – that is, a priori, a series is equally likely to belong to any cluster. Also, recall that the factor innovation variances, λ_k^2, are constant and predetermined.

B.1.2 Notations

- The variance-covariance matrix for each factor k is $\lambda_k^2 \Sigma_k$, where

$$vec(\Sigma_k) = \left(I - \Phi_k^F \otimes \Phi_k^F\right)^{-1} vec\left(u'_{q^f} u_{q^f}\right),$$

and

$$\Phi_k^F = \begin{bmatrix} & \phi_k & \\ I_{q^f-1} & & 0_{q^f-1\times 1} \end{bmatrix}$$

is the companion matrix associated with autoregression (10). u_{q^f} is a $\left(q^f \times 1\right)$ vector with a 1 as the first element and zeros as the rest.

- The variance-covariance matrix for each observation is $\sigma_i^2 \Omega_i$ where

$$vec(\Omega_i) = (I - \mathbb{Z}_i \otimes \mathbb{Z}_i)^{-1} vec\left(u'_{q^\varepsilon} u_{p\varepsilon}\right).$$

$$\mathbb{Z}_i = \begin{bmatrix} & \varphi_i & \\ I_{q^\varepsilon-1} & & 0_{q^\varepsilon-1\times 1} \end{bmatrix}$$

is the companion matrix associated with autoregression (11). u_{q^ε} is a $(q^\varepsilon \times 1)$ vector with a 1 as the first element and zeros as the rest. To quasi-difference the factors following Chib and Greenberg (1994) (Otrok and Whiteman, 1998 as well) I use the matrix defined below;

$$S_i^{-1} = \begin{bmatrix} C_i^{-1} & \mathbf{0} \\ & \Lambda_i^{\varepsilon} \end{bmatrix}.$$

where C_i is the Cholesky matrix of Ω_i, and

$$\Lambda_i^{\varepsilon} = \begin{bmatrix} -\varphi_{iq^{\varepsilon}} & \cdots & -\varphi_{i1} & 1 & 0 & \cdots & \cdots & 0 \\ 0 & -\varphi_{iq^{\varepsilon}} & \cdots & -\varphi_{i1} & 1 & 0 & \cdots & \vdots \\ \vdots & 0 & -\varphi_{iq^{\varepsilon}} & \cdots & -\varphi_{i1} & 1 & 0 & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & -\varphi_{iq^{\varepsilon}} & \cdots & -\varphi_{i1} & 1 \end{bmatrix}.$$

S_i^{-1} is used to quasidifference the data.

- From (10), k^{th} factor measurement error is given by

$$e_{kt} = F_{kt} - \phi_k(L) F_{kt-1}$$

Then, one can stack the factor measurement error as a vector $\widehat{e}_{kq} = \left(e_{k,q^f+1-q}, ..., e_{k,T-q}\right)'$ and define

$$\mathbf{e}_k = \left[\widehat{e}_{k1}, ..., \widehat{e}_{kq^f}\right].$$

- Similarly, From (11) i^{th} idiosyncratic measurement error is given by;

$$\varepsilon_{it} = y_{it} - \alpha_i - \beta_{i,0} F_{0,t} - \sum_{j}^{J} \gamma_{i,j} \beta_{i,j} F_{jt}.$$

Then stack ε_{it} as a vector $\widehat{\varepsilon}_{iq} = (\varepsilon_{i,q^{\varepsilon}+1-q}, ..., \varepsilon_{i,T-q})'$ and define

$$\boldsymbol{\varepsilon}_i = [\widehat{\varepsilon}_{i1}, ..., \widehat{\varepsilon}_{iq^{\varepsilon}}].$$

- Let I be the total number of observations. $\gamma_i = (\gamma_1,\ \gamma_2 ...\ \gamma_J)$ denotes the probability of belonging to clusters from 1 to J for each series i. By assumption only one of γ_i's elements is 1 where all the others are zero.

B.1.3 The Sampler

B.1.3.1 Generating $\boldsymbol{\rho}|\boldsymbol{\Theta}_{-\boldsymbol{\rho}}, \mathbf{F}, \mathbf{Y}$

Given previous draw of factors, the equations in (9) are just i independent regressions with serially correlated errors. Following Chib and Greenberg (1994) we need to account for the serially correlation in the error structure before writing down the distribution of the parameter block. This can be done by building the likelihood for the first q^{ε} observations and continue building the posterior distribution for the rest. To begin define $\mathbf{X}_i^* = [\mathbf{1}_T,\ \mathbf{F}_0]$, where $\mathbf{1}_T$ is a $(T \times 1)$ vector of ones and $\mathbf{F}_0$ is the last draw of the global for series i, respectively and let $\mathbf{Y}_i^* = Y_i - \sum_{j=1}^{J} \gamma_{ij}\beta_{ij}\mathbf{F}_{j,t}$. The following steps lists the process as in Chib and Greenberg (1994)

1. $\widetilde{\mathbf{X}}_{i,1}^* = \begin{bmatrix} 1 & \mathbf{F}_{0,1} \\ ... & ... \\ 1 & \mathbf{F}_{0,q^{\varepsilon}} \end{bmatrix}$ denote the first q^{ε} rows of $\mathbf{X}_i^*$;

2. $\widetilde{\mathbf{Y}}_{i,1}^* = (\mathbf{Y}_{i,1}^*,\ \mathbf{Y}_{i,2}^*, ..., \mathbf{Y}_{i,q^{\varepsilon}}^*)$ denote the first q^{ε} observations of $\mathbf{Y}_i^*$;

3. $\widetilde{\mathbf{X}}_{i,1} = Q_i^{-1}\widetilde{\mathbf{X}}_{i,1}^*$ and $\widetilde{\mathbf{Y}}_{i,1} = Q_i^{-1}\widetilde{\mathbf{Y}}_{i,1}^*$;

4. $\widetilde{\mathbf{X}}_{i,2}$ be a $(T - q^{\varepsilon}) \times 2$ matrix with t^{th} row given by $\varphi_i(L)(\mathbf{X}_{i,t}^*)'$;

5. $\widetilde{\mathbf{Y}}_{i,2}$ be a vector of length $(T - q^{\varepsilon})$ with t^{th} row given by $\varphi_i(L)(\mathbf{Y}_{i,t}^*)'$;

6. Finally, in stack form define $\widetilde{\mathbf{X}}_i = \begin{bmatrix} \widetilde{\mathbf{X}}_{i,1} \\ \widetilde{\mathbf{X}}_{i,2} \end{bmatrix}$ and $\widetilde{\mathbf{Y}}_i = \begin{bmatrix} \widetilde{\mathbf{Y}}_{i,1} \\ \widetilde{\mathbf{Y}}_{i,2} \end{bmatrix}$.

Then for each observation i, $\boldsymbol{\rho}_i = [\alpha_i, \beta_{i0}]'$ is drawn from

$$\boldsymbol{\rho}_i|\boldsymbol{\Theta}_{-\boldsymbol{\rho}_i}, \mathbf{F}, \mathbf{Y} \sim N(\mathbf{r}_i, \mathbf{R}_i),$$

where $\mathbf{R}_i = \left(\mathbf{R}_0^{-1} + \sigma_i^{-2}\widetilde{\mathbf{X}}_i'\widetilde{\mathbf{X}}_i\right)^{-1}$ and $\mathbf{r}_i = \mathbf{R}_i\left(\mathbf{R}_0^{-1}\mathbf{r}_0 + \sigma_i^{-2}\widetilde{\mathbf{X}}_i'\widetilde{\mathbf{Y}}_i\right)$.

B.1.3.2 Generating $\boldsymbol{\sigma}^2|\boldsymbol{\Theta}_{-\boldsymbol{\sigma}^2}, \mathbf{F}, \mathbf{Y}$

σ_i^2 conditional on $\mathbf{Y}$ and $\Theta_{-\sigma_i^2}$, can be drawn from the inverse gamma posterior;

$$\sigma_i^{-2}|\mathbf{Y}, \mathbf{X}, \Theta_{-\sigma_i^2} \sim \boldsymbol{\Gamma}\left(\frac{\nu_0 + T}{2}, \frac{2}{d_0 + d_i'd_i}\right),$$

where $d_i = \widetilde{\mathbf{Y}}_i - \widetilde{\mathbf{X}}_i\rho_i$.

B.1.3.3 Generating $\gamma, \boldsymbol{\beta}|\boldsymbol{\Theta}_{-\gamma,\boldsymbol{\beta}}, \mathbf{F}, \mathbf{Y}$

This step samples the cluster probability and the cluster loadings jointly following FOS (2012). FOS simply utilize an algorithm similar to that of sections 2.5 and 2.6 in Holmes and Held (2006). The joint distribution we are interested in is:

$$p(\boldsymbol{\beta}, \boldsymbol{\gamma}|\boldsymbol{\Theta}, \mathbf{F}) = p(\boldsymbol{\gamma}|\boldsymbol{\Theta}, \mathbf{F})\, p(\boldsymbol{\beta}|\boldsymbol{\Theta}, \boldsymbol{\gamma}, \mathbf{F}).$$

As the closed form for the joint density is not available, this step requires a Metropolis-Hastings sampler within Gibbs draw. Following, Holmes and Held (2006) we can define a joint proposal density, $q(\beta_i^*, \gamma_i^*)$ as;

$$q\left(\beta_i^*, \gamma_i^*\right) = p\left(\beta_i^* | \gamma_i^*, \boldsymbol{\Theta}, \mathbf{F}\right) q\left(\gamma_i^* | \gamma_i\right),$$

where β_i^* and γ_i^* are the candidates and β_i and γ_i are held over from the last draw.

The idea is to draw γ_i^* from a proposal density and to sample β_i^* from its full conditional distribution given this current draw of γ_i^*. The candidates β_i^* and γ_i^* are then accepted with an acceptance probability α. If the candidates are rejected, then the past draws are retained.

The proposal density for γ_i, is assumed to be symmetric in which one draws a random element $\gamma_{i,j}$ and set it equal to 1, while setting all other elements of γ_i to zero. Given this draw of cluster probability, one can sample the candidate β_i^* from the full conditional distribution. In order to compute it, similar to the draw for ρ_i, first define $\overline{\mathbf{X}}_i = \sum_j \gamma_{ij}^* \mathbf{F}_j$ and $\overline{\mathbf{Y}}_i = \left(\mathbf{Y}_i - \alpha_i \mathbf{1}_t - \beta_{0,i} \mathbf{F}_0\right)$, and let;

1. $\overline{\mathbf{X}}_{i,1}^* = \begin{bmatrix} \sum_j \gamma_{i,j}^* \mathbf{F}_{j,1} \\ \ldots \\ \sum_j \gamma_{i,j}^* \mathbf{F}_{j,q^\varepsilon} \end{bmatrix}$ denote the first q^ε rows of $\overline{\mathbf{X}}_i$;

2. $\overline{\mathbf{Y}}_{i,1}^* = (\overline{\mathbf{Y}}_{i,1}, \overline{\mathbf{Y}}_{i,2}, \ldots, \overline{\mathbf{Y}}_{i,q^\varepsilon})$ denote the first q^ε observations of $\overline{\mathbf{Y}}_i$;

3. $\widehat{\mathbf{X}}_{i,1}^* = Q_i^{-1} \overline{\mathbf{X}}_{i,1}^*$ and $\widehat{\mathbf{Y}}_{i,1}^* = Q_i^{-1} \overline{\mathbf{Y}}_{i,1}^*$;

4. $\widehat{\mathbf{X}}_{i,2}^*$ be a vector of length $(T - q^\varepsilon)$ matrix with t^{th} row given by $\varphi_i(L)(\overline{\mathbf{X}}_{i,t})'$;

5. $\widehat{\mathbf{Y}}_{i,2}^*$ be a vector of length $(T - q^\varepsilon)$ with t^{th} row given by $\varphi_i(L) \overline{\mathbf{Y}}_{i,t}$;

6. Finally, in stack form define $\widehat{\mathbf{X}}_i = \begin{bmatrix} \widehat{\mathbf{X}}_{i,1}^* \\ \widehat{\mathbf{X}}_{i,2}^* \end{bmatrix}$ and $\widehat{\mathbf{X}}_i = \begin{bmatrix} \widehat{\mathbf{Y}}_{i,1}^* \\ \widehat{\mathbf{Y}}_{i,2}^* \end{bmatrix}$.

Then,candidate β_i^* can be drawn from the full conditional distribution below:

$$\beta_i | \boldsymbol{\Theta}_{-\boldsymbol{\beta},\boldsymbol{\gamma}}, \gamma_i^*, \mathbf{F}, \mathbf{Y} \sim N\left(\mathbf{b}_i^*, \mathbf{B}_i^*\right), \tag{19}$$

where $\mathbf{B}_i^* = \left(\mathbf{B}_0 + \sigma_i^{-2}\widehat{\mathbf{X}}_i'\widehat{\mathbf{X}}_i\right)^{-1}$, $\mathbf{b}_i^* = \mathbf{B}_i^*\left(\mathbf{B}_0^{-1}\mathbf{b}_0 + \sigma_i^{-2}\widehat{\mathbf{X}}_i'\widehat{\mathbf{Y}}_i\right)$.

The acceptance probability is written in the following form:

$$\alpha = \min\left\{1, \frac{f\left(y|\gamma^*, \beta^*, \Theta_{-\beta,\gamma}, F\right)}{f\left(y|\gamma, \beta, \Theta_{-\beta,\gamma}, F\right)} \frac{p\left(\beta^*\right)}{p\left(\beta\right)} \frac{p\left(\gamma^*\right)}{p\left(\gamma\right)} \frac{q\left(\gamma|\gamma^*\right)}{q\left(\gamma^*|\gamma\right)} \frac{q\left(\beta|\beta^*\right)}{q\left(\beta^*|\beta\right)}\right\}, \tag{20}$$

where the first term is the likelihood; the second term is the prior for γ evaluated at either the candidate or the past draw; the third term is the prior for β; and the last two terms are the probability of a move. This acceptance probability can be simplified further. First off all, $\beta's$ are drawn from the full conditional densities which cancels out the probabilities with $\beta's$ from above. The choice of the symmetric proposal density for γ_n implies that the probability of moving from γ_i^* to γ_i is exactly the same as moving from γ_i to γ_i^*, so that $q\left(\gamma_i^*|\gamma_i\right) = q\left(\gamma_i|\gamma_i^*\right)$. Given also that γ has a uniform prior, which implies $p\left(\gamma^*\right) = p\left(\gamma\right)$, equation(20) reduces to;

$$\alpha = \min\left\{1, \frac{f\left(y|\gamma^*, \beta^*, \Theta, F\right)}{f\left(y|\gamma, \beta, \Theta, F\right)}\right\}.$$

Finally, incorporating the normal likelihoods yields:

$$\alpha_i = \min\left\{1, \frac{|\mathbf{B}_i^*|^{1/2}}{|\mathbf{B}_i|^{1/2}} \frac{\exp\left(\frac{1}{2}\mathbf{b}_i^*\mathbf{B}_i^{*-1}\mathbf{b}_i^*\right)}{\exp\left(\frac{1}{2}\mathbf{b}_i\mathbf{B}_i^{-1}\mathbf{b}_i\right)}\right\}, \tag{21}$$

where $\mathbf{b}_i^*$ and $\mathbf{B}_i^*$ are defined as above and $\mathbf{b}_i$ and $\mathbf{B}_i$ are calculated using the value for γ_i from the past draw. Note that, the draw of the indicator γ_i determines which series enter into the distribution of each group factor.

B.1.3.4 Generating $\phi|\mathbf{\Theta}_{-\phi},\mathbf{F},\mathbf{Y}$

Since the conditional density of $\boldsymbol{\phi}$ has an unknown form, it cannot be sampled directly. I apply Chib and Greenberg (1994) in drawing $\boldsymbol{\phi}=[\boldsymbol{\phi}_0,\boldsymbol{\phi}_1,..,\boldsymbol{\phi}_k]$ conditional on the factors, data, and remaining parameters using a Metropolis-Hastings algorithm. For each iteration, one generates a candidate draw ϕ^* from the proposal distribution below:

$$\boldsymbol{\phi}_k^* \sim N\left(\widehat{\boldsymbol{\phi}}_k,\mathbf{V}_k^{-1}\right),$$

where

$$\mathbf{V}_k=\mathbf{\Phi}_k^{-1}+\lambda_k^2\mathbf{e}_k\mathbf{e}_k{}^{50}$$

and

$$\widehat{\boldsymbol{\phi}}_k=\mathbf{V}_k^{-1}\left(\mathbf{\Phi}_k^{-1}\boldsymbol{\mu}_k+\lambda_k^2\mathbf{e}_k'\widehat{e}_{k0}\right).$$

The candidate ϕ^* is then accepted with a probability that is determined by the likelihood of the data: $\alpha_k=\min\{\widehat{\alpha}_k,1\}$, where

$$\widehat{\alpha}_k=\frac{\Psi\left(\boldsymbol{\phi}_k^*\right)}{\Psi\left(\boldsymbol{\phi}_k^{(n-1)}\right)},$$

and

$$\Psi\left(\boldsymbol{\phi}_k\right)=\left|\Sigma_k\left(\boldsymbol{\phi}_k\right)\right|^{-1/2}\exp\left[-\frac{1}{2\lambda_k^2}\widehat{e}_{k0}'\Sigma_k^{-1}\left(\boldsymbol{\phi}_k\right)\widehat{e}_{k0}\right],$$

with the superscript $n-1$ reflecting the previous iteration. If the draw is less than the acceptance

[50] Refer to the "notations" above for the equation of $\mathbf{e}_k$.

probability, the candidate is accepted. If not, the past draw is retained. Overall, the draw works as follows:

1. First generate the candidate from the proposal density, ϕ^*,

2. Draw from a standard uniform distribution,

3. If the draw is less than the acceptance probability, α_k, set $\phi^{(n)} = \phi^*$

4. Otherwise, retain past draw, $\phi^{(n)} = \phi$.

Generating $\varphi|\boldsymbol{\Theta}_{-\phi}, \mathbf{F}, \mathbf{Y}$

The draw for $\boldsymbol{\varphi}$ follows the same steps as the draw for $\boldsymbol{\phi}$. The autoregression coefficients for the innovation coefficients, $\boldsymbol{\varphi} = [\boldsymbol{\varphi}_1, .., \boldsymbol{\varphi}_i]$, conditional on the factors, data, and remaining parameters are drawn from

$$\boldsymbol{\varphi}_i^* \sim N\left(\boldsymbol{\mu}_i, \boldsymbol{\Upsilon}_i^{-1}\right),$$

where $\boldsymbol{\mu}_i$, $\boldsymbol{\Upsilon}_i$ and the pseudolikelihood $\Psi(\boldsymbol{\varphi}_i)$, follows from above with the necessary change in notation. Here we would have $\Upsilon_i = \bar{\Upsilon}_i + \sigma_i^{-2}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i$, $\boldsymbol{\mu}_i = \boldsymbol{\Upsilon}_i^{-1}\left(\bar{\Upsilon}_i\bar{\boldsymbol{\mu}}_i + \sigma_i^{-2}\boldsymbol{\varepsilon}_i'\widehat{\boldsymbol{\varepsilon}}_{i0}\right)$ and $\boldsymbol{\varphi}_i = |\Omega_i(\boldsymbol{\varphi}_i)|^{-1/2}\exp\left[-\frac{1}{2\sigma_i^2}\widehat{\boldsymbol{\varepsilon}}_{i0}'\Omega_i^{-1}(\boldsymbol{\varphi}_i)\widehat{\boldsymbol{\varepsilon}}_{i0}\right]$.

Generating $\mathbf{F}|\boldsymbol{\Theta}, \mathbf{Y}$

I follow Kim and Nelson (1999) in sampling from the conditional posterior density of the factors given the model's parameters, where the draw of the indicator γ_{ij} determines which series enter into the distribution of each cluster factor. Assume for simplicity that factors and observation errors have the same lag length ($q^f = q^\varepsilon$) and denote it by q. Ley $Y_t = (y_{1,t}, y_{2,t}, \ldots y_{i,t})$, $F_t = (F_{0,t}, F_{1,t} \ldots F_{k,t})$, $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{1,t}, \ldots \varepsilon_{i,t})$, $e_t = (e_{0,t}, e_{1,t}, \ldots e_{k,t})$ and $\epsilon_t = (\epsilon_{0,t}, \epsilon_{1,t}, \ldots \epsilon_{k,t})$ then we can write the state space as by stacking the state variables (factors and observation error terms):

Measurement Equation:

$$[Y_t] = A \ + \ [H] \begin{bmatrix} F_t \\ \varepsilon_t \\ F_{t-1} \\ \varepsilon_{t-1} \\ \vdots \\ F_{t-q} \\ \varepsilon_{t-q} \end{bmatrix}$$

where $A = (\alpha_1 \ , \alpha_2 \ , \ ... \ \alpha_i)'$ and H is $(I \times (K+I)q)$[51] matrix given below:

$$H = \begin{bmatrix} \beta_{1,0} & \gamma_{1,1}\beta_{1,1} & \cdots & \gamma_{i,J} \ \beta_{i,J} & 1 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ \beta_{2,0} & \gamma_{2,1}\beta_{2,1} & \cdots & \gamma_{i,J} \ \beta_{i,J} & 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \beta_{i,0} & \gamma_{i,1}\beta_{i,1} & \cdots & \gamma_{i,J} \ \beta_{i,J} & 0 & 0 & 0 & 1 & 0 & \cdots & 0 \end{bmatrix}$$

with zero variance covariance matrix ($R = 0$), since we stacked all the observation error terms (ε_t) as state variables.

Transition Equation:

[51] $K = 1 + J$ where J is the maximum number of clusters.

$$\begin{bmatrix} F_{t+1} \\ \varepsilon_{t+1} \\ \vdots \\ F_{t+2-q} \\ \varepsilon_{t+2-q} \end{bmatrix} = \mathcal{F} \begin{bmatrix} F_t \\ \varepsilon_t \\ \vdots \\ F_{t-q} \\ \varepsilon_{t-q} \end{bmatrix} + \begin{bmatrix} e_{t+1} \\ \epsilon_{t+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

where $\mathcal{F}$ is $(K+I)q \times (K+I)q$ matrix given by,

$$\mathcal{F} = \begin{bmatrix} diag\left[\phi_{0,1}, \phi_{1,1}, ..., \phi_{J,1}, \psi_{1,1}, ..., \psi_{I,1}\right] & \cdots & diag\left[\phi_{0,q}, \phi_{1,q}, ..., \phi_{J,q}, \psi_{1,q}, ..., \psi_{I,q}\right] \\ 1_{(K+I)(q-1)\times(K+I)(q-1)} & \mathbf{0}_{(K+I)(q-1)\times(K+I)} & \end{bmatrix}$$

with variance-covariance matrix;

$$Q = \begin{bmatrix} diag(\lambda_0^2, \lambda_1^2, ..., \lambda_J^2, \sigma_1^2, ..., \sigma_I^2) & 0_{(K+I)\times(K+I)(q-1)} \\ 0_{(K+I)(q-1)\times(K+I)} & 0_{(K+I)(q-1)\times(K+I)(q-1)} \end{bmatrix}$$

Then the standard Kalman filtering technique can be applied. Let the state vector represented as $\mathbf{S}_t = [F_t \ \varepsilon_t \ ... \ F_{t-q} \ \varepsilon_{t-q}]'$

Given initial values for $\mathbf{S}_{1|0}$ and for the unconditional density of the state vector $P_{1|0}$ the Kalman filter is run from $t = 1$ to $t = T$ following the steps below:

The prediction Step:

$$\begin{aligned} \mathbf{S}_{t|t-1} &= \mathcal{F}\mathbf{S}_{t-1|t-1} \\ P_{t|t-1} &= \mathcal{F}S_{t-1|t-1}\mathcal{F}' + Q \end{aligned}$$

Update:

$$\begin{aligned}\mathbf{S}_{t|t} &= \mathbf{S}_{t|t-1} + P_{t|t-1}H'(H'P_{t|t-1}H + R)^{-1}(Y_t - A - H\mathbf{S}_{t|t-1}) \\ P_{t|t} &= P_{t|t-1} - P_{t|t-1}H'(H'P_{t|t-1}H + R)^{-1}HP_{t|t-1}\end{aligned}$$

Once the Kalman filtering step is over, the factors can be sampled from a Gaussian distribution. However as the Q matrix is singular, an additional step is required to modify these densities as in Kim and Nelson (1999). Their approach is to shrink the size of Q so that it only contains non-zero elements. Here we only want to sample the factors so we reduce the Q matrix to size $K \times K$ and the $\mathcal{F}$ matrix to size $K \times (q(I+K))$. Let $*$'s denote the reduced rank matrices, so we can rewrite:

$$Q^* = diag(\lambda_0^2, \lambda_1^2, ..., \lambda_J^2).$$

and

$$\begin{aligned}\mathcal{F}^* &= \begin{bmatrix}\mathcal{F}_1 & \mathbf{0}_{(I\times I)} & \mathcal{F}_2 & \mathbf{0}_{(I\times I)} & ... & \mathcal{F}_q & \mathbf{0}_{(I\times I)}\end{bmatrix} \text{ where;} \\ \mathcal{F}_n &= diag(\phi_{0,n}, \phi_{1,n}, ..., \phi_{J,n}) \text{ for } n = 1:q.\end{aligned}$$

The last iteration of the Kalman filter provides us with mean, $\mathbf{S}^*_{t|t}$ and variance $P_{t|T}$ which we use to draw the forward period state vector, $\mathbf{S}_{t+1}$ from a Normal distribution. Since, this state vector is full in size we keep its first K elements and denote it as $\mathbf{S}^*_{t+1}$. Now we start to interate backwards to gather the state vector for previous time periods. For $t = T-1, T-2, ...1$, the updating equations are derived as;

$$\begin{aligned}
\mathbf{S}_{t|t,\mathbf{S}^*_{t+1}} &= \mathbf{S}_{t|t} + P_{t|t}\mathcal{F}^{*\prime}(\mathcal{F}^* P_{t|t}\mathcal{F}^{*\prime} + Q^*)^{-1}(\mathbf{S}^*_{t+1} - \mathcal{F}^*\mathbf{S}_{t|t}), \\
P_{t|t,\mathbf{S}^*_{t+1}} &= P_{t|t} - P_{t|t}\mathcal{F}^{*\prime}(\mathcal{F}^* P_{t|t}\mathcal{F}^{*\prime} + Q^*)^{-1}\mathcal{F}^* P_{t|t}.
\end{aligned}$$

Again, we have to keep only the first K elements of $\mathbf{S}_{t|t,\mathbf{S}^*_{t+1}}$ and $P_{t|t,\mathbf{S}^*_{t+1}}$for inference. Let's denote them with $\mathbf{S}^*_{t|t,\mathbf{S}^*_{t+1}}$ and $P^*_{t|t,\mathbf{S}^*_{t+1}}$, respectively. Finally, we can recursively sample the reduced rank state vector, $\mathbf{S}^*_t$, i.e. the factors, form a normal distribution with mean $\mathbf{S}^*_{t|t,\mathbf{S}^*_{t+1}}$ and variance $P^*_{t|t,\mathbf{S}^*_{t+1}}$. This completes the process for the draw of the factors.

B.2 Implementation of Chib's Bayes factor algorithm

The method follows Chib (1995). The marginal likelihood of the model itself is given as:

$$\ln \widehat{m}\left(\mathbf{Y}\right) = \ln f\left(\mathbf{Y}|\mathbf{\Theta}^*\right) + \ln p\left(\mathbf{\Theta}^*\right) - \ln \widehat{p}\left(\mathbf{\Theta}^*|\mathbf{Y}\right),$$

where Θ is the vector of model parameters, $\ln \widehat{m}\left(\mathbf{Y}\right)$ is the log marginal likelihood, $\ln f\left(\mathbf{Y}|\mathbf{\Theta}^*\right)$ is the log likelihood evaluated at a given $\mathbf{\Theta} = \mathbf{\Theta}^*$, $\ln p\left(\mathbf{\Theta}^*\right)$ is the log of the prior evaluated at $\mathbf{\Theta}^*$, and $\ln \widehat{p}\left(\mathbf{\Theta}^*|\mathbf{Y}\right)$ is an approximation of the posterior ordinate. $\mathbf{\Theta}^*$ need only be a high density value of $\mathbf{\Theta}$ (e.g., a modal point). The posterior ordinates can be approximated using the Gibbs output of the full model run. In particular, the posterior ordinate for N sampler blocks that were previously defined is given as;

$$\widehat{p}\left(\mathbf{\Theta}^*|\mathbf{Y}\right) = p\left(\Theta^*_1|\mathbf{Y}\right) \times p\left(\Theta^*_2|\mathbf{Y}, \Theta^*_1\right) \times ... \times p\left(\Theta^*_N|\mathbf{Y}, \Theta^*_1, ..., \Theta^*_{N-1}\right),$$

where a typical term is written as;

$$\widehat{p}\left(\Theta_n^*|\mathbf{Y}\right) = \frac{1}{G}\sum_{g=1}^{G} p\left(\Theta_n^*|\mathbf{Y}, \Theta_1^*, ..., \Theta_{n-1}^*, \Theta_{n+1}^{(g)}, ..., \Theta_N^{(g)}, F_0^{(g)}, F_1^{(g)}, ..., F_J^{(g)}\right).$$

Excluding the latent factors, there are 6 blocks of parameters to determine the posterior ordinates for. Sections below describes each one of them.

B.2.1 Calculation of the posterior ordinate $\widehat{p}\left(\rho^*|\mathbf{Y}\right)$

Define $\mathbf{\Theta} = \left\{\rho, \sigma^2, \varphi, [\gamma, \beta], \phi\right\}$ along with $F_0^{(g)}$, and $F_1^{(g)}, ..., F_J^{(g)}$ where g denotes the number of Gibbs iterations. Let ρ^* be the posterior mode of ρ. Recall that prior of ρ is $N\left(\mathbf{r}_0, \mathbf{R}_0\right)$ and the posterior is $N\left(\mathbf{r}_i, \mathbf{R}_i\right)$ for each i, where and

$$\mathbf{R}_i = \left(\mathbf{R}_0^{-1} + \sigma_i^{-2}\widetilde{\mathbf{X}}_i'\widetilde{\mathbf{X}}_i\right)^{-1},$$

$$\mathbf{r}_i = \mathbf{R}_i\left(\mathbf{R}_0^{-1}\mathbf{r}_0 + \sigma_i^{-2}\widetilde{\mathbf{X}}_i'\widetilde{\mathbf{Y}}_i\right),$$

and $\widetilde{\mathbf{X}}_i$ and $\widetilde{\mathbf{Y}}_i$ are defined appropriately from above. Then, $p\left(\rho^*|\mathbf{Y}\right)$ is approximated by

$$\widehat{p}\left(\rho^*|\mathbf{Y}\right) = \frac{1}{G}\sum_{g=1}^{G} N\left(\rho^*|y, \mathbf{\Theta}^{(g)}, \mathbf{F}^{(g)}\right),$$

where $N(.)$ is the normal pdf with mean and variance defined by the full Gibbs run. As noted in Chib (1995) when calculating the posterior ordinate for the first block we do not need to resample a reduced Gibbs run. Instead the draws from the full Gibbs run should be used to evaluate the following posterior ordinate:

$$\widehat{p}\left(\rho^*|\mathbf{Y}\right) = \frac{1}{G}\sum_{g=1}^{G} N\left(\rho^*, \mathbf{r}_i^{(g)}, \mathbf{R}_i^{(g)}\right), \tag{22}$$

where $\mathbf{r}_i^{(g)}$ and $\mathbf{R}_i^{(g)}$ as defined above are saved from the Full Gibbs run along with the values of $\mathbf{F}^{(g)}$ and $\mathbf{\Theta}^{(g)}$.

B.2.2 Calculation of the posterior ordinate $\widehat{p}\left(\sigma^{2^*}|\mathbf{Y}, \rho^*\right)$

Next, we require $\widehat{p}\left(\sigma^{2^*}|\mathbf{Y}, \mu^*\right)$, which is obtained from an additional (reduced) G runs of the Gibbs sampler holding the previous block fixed at its modal values and sampling all the other model parameters including σ^2 and latent factors,i.e. $\left\{\sigma^{2(g)}, \phi^{(g)}, \varphi^{(g)}, \beta^{(g)}, \gamma^{(g)}, \mathbf{F}^{(g)}\right\}$. Then the posterior estimate is calculated using these draws from the reduced conditional Gibbs run as;

$$\widehat{p}\left(\sigma^{2^*}|\mathbf{Y}, \rho^*\right) = \frac{1}{G}\sum_{g=1}^{G} \Gamma^{-1}\left(\sigma^{2^*}|\mathbf{Y}, \rho^*, \mathbf{\Theta}_{-\rho}^{(g)}, \mathbf{F}^{(g)}\right),$$

where $\Gamma^{-1}(.)$ is the pdf of the inverted gamma distribution. This is operationalized by recalling that σ_i^2 is assumed to have a prior distribution $\sigma_i^2 \sim \Gamma^{-1}\left(\frac{\nu_0}{2}, \frac{\Upsilon_0}{2}\right)$. The posterior distribution from the reduced Gibbs run is saved for each iteration. Then the posterior pdf is evaluated at the modal value of σ^{2^*}. The average across iterations yields the posterior distribution as;

$$\widehat{p}\left(\sigma^{2^*}|\mathbf{Y}, \rho^*\right) = \frac{1}{G}\sum_{g=1}^{G} \Gamma^{-1}\left(\sigma^{2^*}, \frac{\nu_0 + T}{2}, \frac{\Upsilon_0 + \widetilde{\varepsilon}_{iT}^{(g)'}\widetilde{\varepsilon}_{iT}^{(g)}}{2}\right). \tag{23}$$

B.2.3 Calculation of the posterior ordinate $\widehat{p}\left(\phi^*|\mathbf{Y}, \rho^*, \sigma^{2^*}\right)$

Next, we require $\widehat{p}\left(\phi^*|\mathbf{Y}, \rho^*, \sigma^{2^*}\right)$, which – because the parameter is drawn via an MH-in-Gibbs step – is obtained from the method of Chib and Jeliazkov (2001). Their method requires us to save the original draws from the full run and to resample additional G draws of the Gibbs sampler denoted by $\left\{\phi^{(g)},\varphi^{(g)},\beta^{(g)},\gamma^{(g)},\mathbf{F}^{(g)}\right\}$ for the numerator holding the previous blocks fixed at ρ^* and σ^{2^*}. The denominator needs an additional M reduced Gibbs run for $\left\{\varphi^{(g)},\beta^{(g)},\gamma^{(g)},\mathbf{F}^{(g)}\right\}$ holding the aforementioned previous blocks as well as the current block fixed (ϕ^*) at their corresponding modal values. We can then compute

$$\widehat{p}\left(\phi^*|\mathbf{Y},\rho^*,\sigma^{2^*}\right)=\frac{\frac{1}{G}\sum_g\widehat{\alpha}\left(\phi^{(g)},\phi^*|\boldsymbol{\Theta}^{(g)}_{-\rho,\sigma^2},\mathbf{F}^{(g)}\right)\widehat{q}\left(\phi^{(g)},\phi^*|\boldsymbol{\Theta}^{(g)}_{-\rho,\sigma^2},\mathbf{F}^{(g)}\right)}{\frac{1}{M}\sum_j\widehat{\alpha}\left(\phi^*,\phi^{(m)}|\boldsymbol{\Theta}^{(m)}_{-\rho,\sigma^2,\phi},\mathbf{F}^{(m)}\right)}$$

where $\widehat{q}\left(\phi^{(g)},\phi^*\right)=N\left(\widehat{\boldsymbol{\phi}}^*_i,\mathbf{V}_i^{*-1}\right)$ and the acceptance probabilities are defined above. A similar procedure follows for the posterior ordinate of φ.

B.2.4 Calculation of the posterior ordinate $\widehat{p}\left(\beta^*,\gamma^*|\mathbf{Y},\rho^*,\sigma^{2^*},\phi^*\right)$

Next, we require $\widehat{p}\left(\beta^*,\gamma^*|\mathbf{Y},\rho^*,\sigma^{2^*},\phi^*\right)$, which is obtained from both the retained full run and an additional G runs of the Gibbs sampler. Define $\varrho=[\beta,\gamma]$. This step follows from Chib and Jeliazkov (2001) as described in the previous section. The posterior ordinate estimate is calculated with additional G runs for the numerator and additional M runs for the denominator given as below:

$$\widehat{p}\left(\varrho^*|\mathbf{Y},\mu^*,\sigma^{2^*},\phi^*\right)=\frac{\frac{1}{G}\sum_g\alpha\left(\varrho^{(g)},\varrho^*|F^{(g)}\right)q\left(\varrho^{(g)},\varrho^*|\mathbf{F}^{(g)}\right)}{\frac{1}{M}\sum_j\alpha\left(\varrho^*,\varrho^{(m)}|\mathbf{F}^{(m)}\right)},$$

where the proposal density, $q\left(.,.\right)$, and the acceptance probability, $\alpha\left(.,.\right)$ are defined above.[52]

B.2.5 Calculation of the log likelihood evaluated at $\boldsymbol{\Theta}^*$

The log likelihood evaluated at the modal point, $\boldsymbol{\Theta}^*$, can be computed by Monte Carlo integration from the average of the likelihoods for draws of the underlying latent variables:

[52] Note that the notation here is a move from the first to the second.

$$\ln f\left(\mathbf{Y}|\Theta^*\right) = \frac{1}{G}\sum_{g=1}^{G}\ln f\left(\mathbf{Y}|\Theta^*, \mathbf{F}^{(g)}\right). \tag{24}$$

To compute this, we would set the model parameters at the mode and use factors sampled from the full Gibbs run to compute the likelihood at each point. The log-likelihood would then be the average of these likelihoods across iterations.

B.2.6 Calculation of the prior evaluated at $\mathbf{\Theta}^*$

The second term in (12) represents the prior distributions evaluated at their modal values and can be evaluated as

$$\ln p\left(\mathbf{\Theta}^*|y\right) = \ln p\left(\Theta_1^*\right) + \ln p\left(\Theta_2^*\right) + ... + \ln p\left(\Theta_i^*\right).$$

B.3 Application of Bayesian Linear Regression

We are interested in approximating equation (14). To do so, we can save every 50th draw from the full posterior distribution of the model factors, and run Bayesian linear regression on each of the saved factor draws. Ξ_t represents the set of the variables we want to test on the factors $F_{k,t}$. Then we can estimate the linear regression of the form

$$F_{k,t} = \varpi_k \Xi_t + \upsilon_k$$

where the error term υ is assumed to be normally distributed with mean zero and variance $\oplus_k^2$. The estimation is a simple Gibbs application with two sampler blocks of parameters, namely ϖ and $\oplus^2$. Let the prior distributions for the coefficients and the variance be represented with $N(a, b)$ and $IG\left(\frac{c}{2}, \frac{d}{2}\right)$ respectively $(a = 0, b = 2, c = 6, d = 0.1)$.

B.3.1 Generating $\boldsymbol{\varpi}|\oplus^2\mathbf{F},\Xi$

The conditional distribution can be drawn from

$$\varpi_k|\oplus^2,\mathbf{F},\Xi{\sim}N\left(\mathbf{A}_k,\mathbf{B}_k\right),$$

where $\mathbf{A}_k=\left(\mathbf{b}^{-1}+\oplus_k^{-2}\Xi'\Xi\right)^{-1}$ and $\mathbf{B}_k=\mathbf{A}_k\left(\mathbf{b}^{-1}\mathbf{a}+\oplus_k^{-2}\Xi' F_k\right)$.

B.3.2 Generating $\oplus^2|\boldsymbol{\varpi},\mathbf{F},\Xi$

$\oplus^{-2}$ conditional on $\mathbf{F},\boldsymbol{\Xi}$ and $\boldsymbol{\varpi}$, can be drawn from the gamma posterior;

$$\oplus_k^{-2}|\boldsymbol{\varpi},\mathbf{F},\boldsymbol{\Xi}\sim\boldsymbol{\Gamma}\left(\frac{c+T}{2},\frac{2}{d+D_k'D_k}\right),$$

where $D_k=\mathbf{F}_k-\varpi_k\Xi_t$.

Recall that for each factor k, we saved every 50^{th} draw. We apply the steps above to each saved draw of $F_{k,t}$ for 1000 iterations while burning in the first 500. This gives us many posterior distributions for each block ϖ_k and $\oplus_k^{-2}$. Then the posterior distributions of these parameter blocks are pooled together from which we can make inferences. From these pooled posterior distributions I report the mean of $\boldsymbol{\varpi}$ along with the Bayesian confidence intervals. The confidence interval is the 5^{th} and 95^{th} percentile interval endpoints of the pooled distribution.

B.4 Data

The list of primary commodities and their explanations are directly taken from IFS database.

Wheat: United States, No.1 Hard Red Winter, ordinary protein, FOB Gulf of Mexico, US$ per metric tonne.

Maize (corn): United States. No.2 Yellow, FOB Gulf of Mexico, U.S. price, US$ per metric tonne.

Rice: Thailand, 5 percent broken milled white rice, Thailand nominal price quote, US$ per metric tonne.

Barley: Canada, no.1 Western Barley, spot price, US$ per metric tonne.

Soybean Meal: United States, Chicago Soybean Meal Futures (first contract forward) Minimum 48 percent protein, US$ per metric tonne.

Soybean Oil: United States, Chicago Soybean Oil Futures (first contract forward) exchange approved grades, US$ per metric tonne.

Soybeans: United States., Chicago Soybean futures contract (first contract forward) No. 2 yellow and par, US$ per metric tonne.

Fishmeal: Peru, Fish meal/pellets 65% protein, CIF, US$ per metric tonne.

Sunflower oil: United Kingdom, US export price from Gulf of Mexico, US$ per metric tonne.

Olive Oil: United Kingdom, extra virgin less than 1% free fatty acid, ex-tanker price U.K., US$ per metric tonne.

Palm oil: Malaysia, Palm Oil Futures (first contract forward) 4-5 percent FFA, US$ per metric tonne.

Rapeseed (referred as Canola) oil: United Kingdom, crude, fob Rotterdam, US$ per metric tonne

Groundnuts (peanuts): Nigeria, 40/50 (40 to 50 count per ounce), US$ per metric tonne.

Beef: Australia and New Zealand, 85% lean fores, CIF U.S. import price, US cents per pound.

Lamb: New Zealand, frozen carcass Smithfield London, US cents per pound.

Swine (pork): United States, 51-52% lean Hogs, US cents per pound.

Poultry (chicken): United States, Whole bird spot price, Ready-to-cook, whole, iced, Georgia docks, US cents per pound.

Fish (salmon): Norway, Farm Bred Norwegian Salmon, export price, US$ per kilogram.

Shrimp: United States, No.1 shell-on headless, 26-30 count per pound, Mexican origina, New York port, US cents per pound.

Sugar: World, Free Market, Coffee Sugar and Cocoa Exchange (CSCE) contract no.11 nearest future position, US cents per pound.

Sugar: United States, U.S. import price, contract no.14 nearest futures position, US cents per pound (Footnote: No. 14 revised to No. 16).

Oranges: France, miscellaneous oranges CIF French import price, US$ per metric tonne.

Bananas: Central America and Ecuador, FOB U.S. Ports, US$ per metric tonne.

Coffee: Africa not specified, Robusta, International Coffee Organization New York cash price, ex-dock New York, US cents per pound.

Coffee, Other Mild Arabicas, International Coffee Organization New York cash price, ex-dock New York, US cents per pound.

Cocoa beans: Ghana, International Cocoa Organization cash price, CIF US and European ports, US$ per metric tonne.

Tea: Mombasa, Kenya, Auction Price, US cents per kilogram, From July 1998,Kenya auctions, Best Pekoe Fannings. Prior, London auctions, c.i.f. U.K. warehouses.

Hard Logs: Malaysia, Best quality Malaysian meranti, import price Japan, US$ per cubic meter.

Soft Logs: United States, Average Export price from the U.S. for Douglas Fir, US$ per cubic meter.

Hard Sawnwood: Malaysia, Dark Red Meranti, select and better quality, C&F U.K port, US$ per cubic meter.

Soft Sawnwood: United States, average export price of Douglas Fir, U.S. Price, US$ per cubic meter.

Cotton: United States, Cotton Outlook 'A Index', Middling 1-3/32 inch staple, CIF Liverpool, US cents per pound.

Wool coarse: United Kingdom, 23 micron, Australian Wool Exchange spot quote, US cents per kilogram.

Wool fine: United Kingdom, 19 micron, Australian Wool Exchange spot quote, US cents per kilogram.

Rubber: Malaysia, No.1 Smoked Sheet, Singapore Commodity Exchange, 1st contract, US cents per pound.

Hides: United States, Heavy native steers, over 53 pounds, wholesale dealer's price, Chicago, fob Shipping Point, US cents per pound.

Copper: United Kingdom, grade A cathode, LME spot price, CIF European ports, US$ per metric tonne.

Aluminum: Canada, 99.5% minimum purity, LME spot price, US$ per metric tonne.

Iron Ore: China import 62% FE spot (CFR Tianjin port), US cents per dry metric tonne unit.

Tin: United Kingdom, standard grade, LME spot price, US$ per metric tonne.

Nickel: United Kingdom, melting grade, LME spot price, CIF European ports, US$ per metric tonne.

Zinc: United Kingdom, high grade 98% pure, US$ per metric tonne.

Lead: United Kingdom, 99.97% pure, LME spot price, CIF European Ports, US$ per metric tonne.

Uranium: World, NUEXCO, Restricted Price, Nuexco exchange spot, US$ per pound.

Crude Oil: Arab Emirates, Dubai, medium, Fateh 32 API, fob DubaiCrude Oil (petroleum), Dubai Fateh Fateh 32 API, US$ per barrel.

REFERENCES

Ahmed, Shaghil; Ickes, Barry W.; Wang, Ping; and Yoo, Byung Sam. "International Business Cycles." American Economic Review, June 1993, 83(3), pp. 335-59.

Akram Q. F., "Commodity prices, interest rates and the dollar", Energy Economics, Volume 31, Issue 6, November 2009, Pages 838-851

Andrews, D.W.K., "Exactly Median-Unbiased Estimation of First Order Autoregressive/Unit Root Models," Econometrica 61: (1993), 139-165.

Andrews, D.W.K., and H.-Y. Chen, "Approximately Median-Unbiased Estimation of Autoregressive Models," Journal of Business & Economic Statistics, Vol. 12, No. 2 (Apr., 1994), pp. 187-204

Baffes, J., "More on the Energy/Non-Energy Commodity Price Link" World Bank Policy Research Working Paper Series (2009)., No. 4982

Baffes, J. and T. Haniotis, "Placing the 2006/08 Commodity Price Boom into Perspective", World Bank Policy Research Working Paper 5371 (2010)

Bai, J. "Inference on factor models of large dimensions." Econometrica 71(1), 2003, pp. 135.172.

Bai, J. and Ng, S. "Determining the number of factors in approximate factor models." Econometrica 70(1), 2002, pp. 191.221.

Baxter, Marianne & Kouparitsas, Michael A., 2005. "Determinants of business cycle comovement: a robust analysis." Journal of Monetary Economics, Elsevier, vol. 52(1), pages 113-157, January

Baxter, Marianne & Kouparitsas, Michael A., 2005. "Determinants of business cycle comovement: a robust analysis." Journal of Monetary Economics, Elsevier, vol. 52(1), pages 113-157, January

Bernanke, B.S., Boivin, J., P. Eliasz, "Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach", Quarterly Journal of Economics 120, (2005) , 387–422

Borensztein E. and C. M. Reinhart, "The Macroeconomic Determinants of Commodity Prices", IMF Staff Papers, Vol. 41 (1994), No. 2, 236-258.

Boivin J. and S. Ng. "Are more data always better for factor analysis?" Journal of Econometrics 132, 2006.

Byrne, J. P., G. Fazio, and N. M. Fiess, "Primary Commodity Prices: Co-Movements, Common Factors and Fundamentals", World Bank Policy Research Working Paper Series (2011), Vol. , pp.

Carter, C.K. and Kohn, R. “On Gibbs Sampling for State Space Models.” Biometrika, August 1994, 81(3), pp. 541-53.

Carter, C., G. Rausser and A. Smith, "Commodity Booms and Busts", Annual Review of Resource Economics (2011).

Carvalho, Carlos M.; Lopes, Hedibert F.; and Aguilar, Omar. “Dynamic Stock Selection Strategies: A Structural Factor Model Framework.” In Bayesian Statistics 9, J.M. Bernardo et al. (eds.). Oxford: Oxford University Press, 2010.

Calvo, G. "Exploding Commodity Prices, Lax Monetary Policy, and Sovereign Wealth Funds".Vox EU. (2008)

Casella, George and George, Edward I. “Explaining the Gibbs Sampler.” American Statistician, August 1992, 46(3), pp. 167-74.

Cashin, P., C. J. McDermott, and A. Scott, "The Myth of Co-Moving Commodity Prices", Bank of New Zealand Discussion Paper (1999) No. G99/9.

Cashin, P., C. J. McDermott and A. Scott, "Booms and Slumps in World Commodity Prices", Journal of Development Economics (2002), vol. 69, pp. 277-296

Cashin P., H. Liang, C. and J. McDermott, "How Persistent Are Shocks to World Commodity Prices?", IMF Staff Papers, Vol. 47, No. 2 (2000), pp. 177-217

Chen, Shu-Ling, J. D. Jackson, H. Kim, and P. Resiandini, "What Drives Commodity Prices?", (2010) No auwp2010-05, Auburn Economics Working Paper Series, Department of Economics, Auburn University.

Chen, Shiu-Sheng, and C. Engel, “Does “Aggregation Bias” Explain the PPP Puzzle?” Pacific Economic Review 10, (February 2005), 49-7.

Chib, Siddhartha. “Bayes Estimation of Regressions with Autoregressive Errors: A Gibbs Sampling Approach.” Journal of Econometrics, August 1993, 58(3), pp. 275-94.

Chib, Siddhartha. “Marginal Likelihood from the Gibbs Output.” Journal of the American Statistical Association, December 1995, 90(432), pp. 1313-21.

Chib, Siddhartha and Greenberg, Edward. “Bayes Inference in Regression Models with ARMA (p,q) Errors.” Journal of Econometrics, September-October 1994, 64(1-2), pp. 183-206.

Chib, Siddhartha and Jeliazkov, Ivan. “Marginal Likelihood from the Metropolis-Hastings Output.” Journal of the American Statistical Association, March 2001, 96(453), pp. 270-81.

Chib, Siddhartha and Jeliazkov, Ivan. “Marginal Likelihood from the Metropolis-Hastings Output.” Journal of the American Statistical Association, March 2001, 96(453), pp. 270-81.

Clark, Todd E. “Employment Fluctuations in U.S. Regions and Industries: The Roles of National, Region-Specific, and Industry-Specific Shocks.” Journal of Labor Economics, January 1998, 16(1), pp. 202-29.

Clark, Todd E. and Shin, Kwanho. “The Sources of Fluctuations within and across Countries.” In Intranational Macroeconomics, G.D. Hess and E. van Wincoop, eds. Cambridge: Cambridge University Press, 2000, pp. 189-217.

Clark, Todd and van Wincoop, Eric. "Borders and Business Cycles." Journal of International Economics, 55, 2001, pp. 59-85.

Cuddington, J. T. and D. Jerrett, "Super Cycles in Real Metals Prices?" , IMF Staff Papers (2008), Vol. 55, Issue 4, pp. 541-565.

Devroye, Luc. Non-Uniform Random Variate Generation. New York: Springer-Verlag, 1986.

Djankov; La Porta; Lopez-de-Silanes; and Shleifer. “Courts.” Quarterly Journal of Economics, May, 2003

Engle, Robert and Sharon Kozicki (1993), "Testing for Common Features", Journal of Business and Economic Statistics, 11, 369-380.

Francis, N. R., M. T. Owyang and O. Savascin, "An Endogenously Clustered Factor Approach to International Business Cycles" St. Louis Fed Working paper series (April 2012)

Frankel, J.A. "Why are Commodity Prices so High? Don’t Forget Low Interest Rates", Financial Times, (2005). 4/15/05

Frankel, J.A. "The Effect of Monetary Policy on Real Commodity Prices", Asset Prices and Monetary Policy (2008), University of Chicago Press

Frankel J. A. and A. K. Rose, "Determinants of Agricultural and Mineral Commodity Prices", Working Paper, Kennedy School of Government, Harvard University, (2009).

Frühwirth-Schnatter, S. and S. Kaufmann, “Model-Based Clustering of Multiple Times Series.” Journal of Business and Economic Statistics, January 2008, 26(1), pp. 78-89.

Frühwirth-Schnatter, Sylvia and Lopes, Hedibert Freitas. “Parsimonious Bayesian Factor Analysis when the Number of Factors Is Unknown.” Unpublished, 2009.

Foerster, A. T., P. G. Sarte, and M. W. Watson, “Sectoral versus aggregate shocks: A structural factor analysis of industrial production.” Journal of Political Economy (2011), 119,1–38

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. “The generalized factor model: Identification and estimation,” Review of Economics and Statistics 82, 2000, pp. 540.554.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. “The generalized factor model: One-sided estimation and forecasting,” WP (2005), forthcoming, Journal of the American Statistical Association.

Gelfand, Alan E. and Smith, Adrian F.M. “Sampling-Based Approaches to Calculating Marginal Densities.” Journal of the American Statistical Association, June 1990, 85(410), pp. 398-409.

Ghosh, J. and Dunson, D.B. “Bayesian model selection in factor analytic models.” In Random Effect and Latent Variable Model Selection, 2008, ed. D.B. Dunson. John Wiley & Sons.

Hallin M. and R. Liska. “Dynamic factors in the Presence of Block Structure,” EUI Working Papers, 2008, ECO 2008/22.

Hamilton, J. D. "Causes and Consequences of the Oil Shock of 2007-2008", NBER Working Paper Series (2009), 15002

Hamilton, J.D. "Time Series Analysis", Princeton University Press (1994).

Hamilton, James D. and Owyang, Michael T. “The Propagation of Regional Recessions.” WP (2009), forthcoming, Review of Economics and Statistics.

Holmes, Chris C. and Held, Leonhard. “Bayesian Auxiliary Variable Models for Binary and Multinomial Regression.” Bayesian Analysis, 2006, 1(1), pp. 145-168.

Hochman, G., D. Rajagopal and D. Zilberman, "Are Biofuels the Culprit? OPEC, Food, and Fuel". American Economic Review, Vol.100, No.2, (May 2010), pp. 183-18

Heston, Alan; Summers, Robert; and Aten, Bettina. Penn World Table Version 6.3, Center for International Comparisons of Production, Income, and Prices at the University of Pennsylvania, August 2009.

Holmes, Chris C. and Held, Leonhard. “Bayesian Auxiliary Variable Models for Binary and Multinomial Regression.” Bayesian Analysis, 2006, 1(1), pp. 145-168.

Imbs, J., H. Mumtaz, M. O. Ravn, and H. Rey, “PPP Strikes Back: Aggregation and the Real Exchange Rate,” NBER Working Paper 9372, (2003).

Kaufmann, Silvia. “Dating and Forecasting Turning Points by Bayesian Clustering with Dynamic Structure: A Suggestion with an Application to Austrian Data.” Journal of Applied Econometrics, March 2010, 25(2), pp. 309-44.

Kilian, Lutz, “Small-Sample Confidence Intervals for Impulse Response Functions,” Review of Economics and Statistics LXXX (II) (1998), 218-30.

Kim, Chang-Jin and Charles Nelson, “Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle”, Review of Econ and Statistics, 1999, 81(4), 608-616.

Kim, Chang-Jin and Charles Nelson, "State-Space Models with Regime Switching: Classical and Gibbs Sampling Approaches with Applications", Cambridge: MIT Press, 1999.

Kose, M. Ayhan; Otrok, Christopher; and Whiteman, Charles H. “International Business Cycles: World, Region, and Country-Specific Factors.” American Economic Review, September 2003, 93(4), pp. 1216-39.

Kose, M. Ayhan; Otrok, Christopher; and Whiteman, Charles H. “Understanding the Evolution of World Business Cycles.” Journal of International Economics, May 2008, 75(1), pp. 110-30.

Krugman, P. "The Oil Nonbubble". The New York Times. (2008) 12 May.

La Porta; Lopez-de-Silanes; Shleifer; and Vishny. “The Quality of Government.” Journal of Law, Economics and Organization, 1999.

Lescaroux, F. "On the excess co-movement of commodity prices-A note about the role of fundamental factors in short-run dynamics", Energy Policy, Volume 37, Issue 10 (2009), pp. 3906-3913

Lu, Y. and S. N. Neftci, "Financial Instruments to Hedge Commodity Price Risk for Developing Countries", International Monetary Fund Working Papers (2008), 08/6.

Lombardi, Marco J., Osbat, Chiara and Schnatz, Bernd, "Global Commodity Cycles and Linkages: A FAVAR Approach", European Central Bank Working Paper (2010), No. 1170.

Martin, A. Bayesian Analysis. In J. Box-Steffensmeier & D. Collier (Eds.), The Oxford Handbook of Political Methodology, Oxford University Press, (2008).

McKinnon, Ronald I. “Currency Substitution and Instability in the World Dollar Standard.” American Economic Review, June 1982, 72(3), pp. 320-33.

Moench, Emanuel; Ng, Serena; and Potter, Simon. “Dynamic Hierarchical Factor Models.” Federal Reserve Bank of New York Staff Report No. 412, December 2009.

Mueller, S. A., J. E. Anderson and T. J. Wallington, "Impact of biofuel production and other supply and demand factors on food price increases in 2008", Biomass and Bioenergy (2011), Volume 35, Issue 5, Pages 1623-1632

Neely, C. J. and D. Rapach, "International Comovements in Inflation Rates and Country Characteristics" (June 13, 2011). Federal Reserve Bank of St. Louis Working Paper No. 2008-025

Norrbin, Stefan C. and Schlagenhauf, Don E. “The Role of International Factors in the Business Cycle: A Multi-Country Study.” Journal of International Economics, February 1996, 40(1-2), pp. 85-104.

Onatski A., “Asymptotics of the Principal Components Estimator of Large Factor Model with Weak Factors and i.i.d. Gaussian Noise." 2007, mimeo, Columbia University

Otrok, C. and C. H. Whiteman, "Bayesian Leading Indicators: Measuring and Predicting Economic Conditions in Iowa" International Economic Review, November (1998), 39(4), pp. 997-1014.

Paap, R., "What are the advantages of MCMC based inference in latent variable models?" Statistica Neerlandica 56, (2002), 2-22.

Pesaran, M. H. and Z. Zhao, "Bias Reduction in Estimating Long-run Relationships from Dynamic Heterogeneous Panels," in Analysis of Panels and Limited Dependent Variables, Cambridge University Press, (1999), chapter 12, 297-32

Phillips, Peter and Donggyu Sul, "Bias in dynamic panel estimation with fixed effects, incidental trends and cross section dependence," Journal of Econometrics, Volume 137, Issue 1, (March 2007), 162-188.

Pindyck, R. S. and J. J. Rotemberg. "The excess co-movement of commodity prices", Economic Journal, Vol. 100, (1990) pp. 1173-89.

Ponomareva, Natalia and Katayama, Hajime. "Does the Version of the Penn World Tables Matter? An Analysis of the Relationship Between Growth and Volatility." Canadian Journal of Economics, February 2010, 43(1), pp. 152-79.

Sargent, T.J. and C.A. Sims, "Business Cycle Modeling Without Pretending to Have Too Much A Priori Economic Theory," in Christopher A. Sims et al., eds., New Methods in Business Cycle Research, (1977), pp. 45-108).

Stock, J. H. and Watson, M. W. "Forecasting using principal components from a large number of predictors." Journal of the American Statistical Association 97, 2002a, pp 1167-79.

Stock, J. H. and Watson, M. W. "Macroeconomic forecasting using diffusion indices." Journal of Business and Economic Statistics 20(2), 2002b, pp. 147-62.

Svensson, L.E.O., "The effect of monetary policy on real commodity prices: Comment", In Asset Prices and Monetary Policy, Ed. John Y. Campbell, NBER Working Paper (2008) 12713.

Tanner, Martin A. and Wong, Wing Hung. "The Calculation of Posterior Distributions by Data Augmentation." Journal of the American Statistical Association, June 1987, 82(398), pp. 528-40.

Trostle R. "Global Agricultural supply and demand: factors contributing to the recent increase in food commodity prices", USDA Economic Research Service (July 2008), Report WRS-0801

Troughton, P. T. and S. J. Godsill. "Bayesian model selection for time series using Markov chain Monte Carlo. In Proc. IEEE International Conference on Acoustics, Speech and Signal Processing", April 1997

Vansteenkiste, I., "How important are common factors in driving non-fuel commodity prices? A dynamic factor analysis", European Central Bank Working Paper (2009), no. 1072.

Western, B., “Causal Heterogeneity in Comparative Research: A Bayesian Hierarchical Modelling Approach,” American Journal of Political Science, (1998), 42:1233–1259.

Wolf, M. "Life in a tough world of high commodity prices", Financial Times, (2008), March 4.