Research Division
Federal Reserve Bank of St. Louis
Working Paper Series

An Endogenously Clustered Factor Approach to International Business Cycles

Neville Francis
Michael T. Owyang
And
Özge Savascin

Working Paper 2012-014A
http://research.stlouisfed.org/wp/2012/2012-014.pdf

April 2012

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
P.O. Box 442
St. Louis, MO 63166

The views expressed are those of the individual authors and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors.

Federal Reserve Bank of St. Louis Working Papers are preliminary materials circulated to stimulate discussion and critical comment. References in publications to Federal Reserve Bank of St. Louis Working Papers (other than an acknowledgment that the writer has had access to unpublished material) should be cleared with the author or authors.

An Endogenously Clustered Factor Approach to International Business Cycles*

Neville Francis
Department of Economics
University of North Carolina, Chapel Hill

Michael T. Owyang†
Research Division
Federal Reserve Bank of St. Louis

Özge Savaşçın
Department of Economics
University of North Carolina, Chapel Hill

Keywords: factor model, block factor, logistic prior, cluster analysis, hierarchical prior

April 10, 2012

Abstract

Factor models have become useful tools for studying international business cycles. Block factor models [e.g., Kose, Otrok, and Whiteman (2003)] 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 Monte Carlo experiments, we show that even small misspecification can lead to substantial declines in fit. 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 us to incorporate series-level covariates that may influence and explain how the series are grouped. Using similar international business cycle data as Kose, Otrok, and Whiteman, we find our country clusters differ in important ways from 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.

[JEL: C33; C52; E32; F44]

*Kristie M. Engemann, Charles Gascon, and Kate Vermann provided research assistance. The authors benefited from comments from or discussions with Laura Coroneo, Jim Hamilton, Chris Otrok, Jeremy Piger, and Dave Rapach and comments from seminar participants at Emory University and the University of California at Davis. Conference participants at the Duke University Special on Applied Macro-Time Series and Macro Workshop and the Applied Econometrics Workshop at the Federal Reserve Bank of St. Louis supplied useful suggestions. The views expressed here are the authors' alone and do not reflect the opinions of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

†corresponding author: P.O. Box 442, St. Louis, MO 63166. owyang@stls.frb.org

1 Introduction

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 the countries that 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.[2] 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.[3] Finally, correlation between macroeconomic aggregates across countries could be due to unobservable innovations – e.g., common international shocks or country-specific shocks having spillover effects. Using structural vector autoregressions, Ahmed, Ickes, Wang, and Yoo (1993) concludes that spillovers from a 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 approach leaves the country combinations unrestricted (i.e., any two countries' cycles can be correlated), whereas the latter explicitly excludes this. The choice of the approach taken can depend on both the question to be answered and the econometric techniques used to

[1] The relationship among 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] See Baxter and Kouparitsas (2005) for a list of other potential determinants of business cycle comovements across countries.

[3] 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 geographic borders in the synchronization of business cycles between U.S. Census regions and across European countries, Clark and van Wincoop (2001) found limited roles for both monetary and fiscal policies.

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.[4] The second approach might define a common cycle via a factor model, where the factor loadings reflect the degree of correlation among 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.[5] This block structure provides a straightforward interpretation that may be lacking in standard factor models. Countries within a block have cycles that are correlated through a regional factor, whereas 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 zero. Even in that case, however, the factors 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 a small number 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 significant *ex ante* assumptions must be made 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 be useful in determining important features that any successful model of the international business cycle should possess. For example, if we find that common language is a better determinant of cross-correlation than physical distance, models of trade may consider common language rather than geography as the determinant of iceberg costs.

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

[5] See also Boivin and Ng (2006); Onatski (2012); and Hallin and Liska (2008).

The model has the block factor structure with an additional membership indicator determining to which block a country belongs. 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.[6] In the simplest execution of 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 compared with 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 (1) incorporate country-specific characteristics (e.g., location, industrialization, trade patterns) and (2) test competing hypotheses about which influences determine the countries that 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. When 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 dramatic deviations from the true model and substantial reductions in fit.

Our empirical application extends KOW's study of cross-country correlations. Using annual gross domestic product (GDP) growth rates for 60 countries, we find that although some regional/geographic correlation does exist, 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 many of the industrialized nations: Japan and most of Europe, excluding the U.K. and Denmark. 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 substantially higher contribution of the cluster (or regional)

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

factor to the overall volatility of output.

The balance of the paper is organized as follows: Section 2 presents the endogenous clustered factor model. Section 3 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 3.2 presents some Monte Carlo evidence showing how well our algorithm identifies the clusters and the consequences of exogenously misidentifying them. Section 4 extends the model and the sampler with a multinomial logistic prior. Section 5 presents results from the model with international business cycle data. Section 6 summarizes and concludes.

2 Empirical Model

Suppose that we have 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 that affect all series and those that affect only a few series. We refer to the former as global factors and 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.[7] 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}:

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

$i = 1, \ldots, M,\ t = 1, \ldots, T,\ n = 1, \ldots, N$ 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.[8] 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

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

[8] 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)].

the restriction that the countries on the same continent comove; Moench, Ng, and Potter impose that within-sector data comove. While geographic proximity or industrial similarity may be a reason for the comovement between two countries, 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, $\gamma_{ni} = \{0, 1\}$, that signifies whether series n belongs to cluster i, retaining the restriction that a series can only belong 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 clusters comove only through the global factor, while series belonging to 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.

3 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) is obtained by computing the marginal likelihoods for models with different numbers of clusters [see also Ghosh and Dunson (2008)].[9]

3.1 The Sampler

The sampler is an MCMC algorithm that 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 AR parameters, $\boldsymbol{\psi}$; (4) the factors, $\mathbf{F}$; and (5) the set of factor AR parameters, $\boldsymbol{\phi}$. After initializing the sampler, the posterior distributions are computed with 10,000 iterations after 30,000 iterations are discarded for convergence.

3.1.1 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 – that is, *a priori*, a series is equally likely to belong to any cluster. In section 4, we modify the sampler to incorporate

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

country-specific characteristics into the cluster determination through a logistic heirarchical 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 AR parameters are straightforward and included in Appendix A.

3.1.2 Preliminaries

Before discussing conditional distributions for each block, it is 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

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

which 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 $\Sigma_n = \left(\mathbf{I} - \Psi_i \otimes \Psi_i\right)^{-1} vec\left(u'_{p_\varepsilon} u_{p_\varepsilon}\right)$ 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.

3.1.3 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 $\mathbf{Y}_n$.[10] Then, the candidate β_n^* is drawn from

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

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 $\pi\left(\beta_n^* | \gamma_n^*, \boldsymbol{\Theta}_{-\beta,\gamma}, \mathbf{F}, \mathbf{Y}\right)$, the value of β_n^* does not appear in the acceptance probability.[11] In this case, for each n, the acceptance probability is

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

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 1 (setting all other elements equal to 0). The choice of the symmetric proposal makes the last term in (5)

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

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

identically 1.[12]

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

The factors are drawn recursively from the smoothed Kalman update densities using the techniques 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 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 that 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 Appendix A.

3.2 The Effect of Misspecification

Allowing the data to determine the clusters rather than setting them in advance highlights a trade-off 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 the extent to which the clusters must be misspecified to outweigh the uncertainty of estimating them.

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, a 1.7 percent misspecification refers to one series allocated to the wrong cluster with all other series correctly specified. We then estimate

[12] 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.

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.[13]

Table 2 reports the results of the MC experiments. As expected, lower 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 note that, in these experiments, we give the best chance to pre-specification of the clusters by correctly setting the true number of clusters – that is, the only source of potential misspecification is incorrectly assigning a series n to the wrong cluster.

4 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 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.

[13] 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.

4.1 Adding a Prior for Cluster Membership

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

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

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 considered a draw from a truncated logistic distribution. We follow Holmes and Held (2006) by defining a new latent variable, χ_{ni}, that allows 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:

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

If $\chi_{ni}\sim KS$ and $o_{ni}\sim N\left(0,1\right)$, 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.[14] The cluster probabilities can be rewritten in terms of the new latent variables as

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

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

[14] See Devroye (1986).

4.2 Augmenting the Sampler

The sampler outlined in Section 3 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.

4.2.1 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 i 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

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

4.3 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 such 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).

5 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 that have been suggested as affecting trade between countries. By doing this, we can assess the sources of business cycle comovements.

5.1 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) from the 6.3 version of the Penn World Tables (PWT) [Heston, Summers, and Aten (2009)].[15] To maintain comparibility, we choose the same 60 countries located in 7 regional blocks as in KOW.[16]

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 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 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

[15] 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.

[16] To increase the 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 A.1 in the data appendix shows the 60 countries in the estimation along with the regional groupings imposed in KOW.

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. The 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, La Porta, Lopez-de-Silanes, and Shleifer (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 setups.

The ethnolinguistic index in (4) is from La Porta, Lopez-de-Silanes, Shleifer, and Vishny (1999) and measures the degree of language diversity, the probability that two randomly selected individuals in a given country speak different languages, do not speak the official language, or do not speak the most widely used language.

Finally, Baxter and Kouparitsas (2005) 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.[17] Baxter and Kouparitsas find that industrialized nations have dispersions similar to the rest of the world (the average country) for all three indices, whereas 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 trade relations for developing nation are spread more

[17] Appendix B provides more details on the construction of these indices.

evenly across developed and developing nations. By including these indices, we are allowing for the possibility that countries form clusters based 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 the types of goods traded and the trading partners).

5.2 Results

We first determine the optimal number of country-level factors which, for simplicity, we compute with a flat heirarchical 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. The moels with two and six clusters have the next-highest marginal likelihood; however, either alternative requires more than 100 times higher prior likelihood to be preferred. The model with seven clusters – the specification that nests the one estimated by KOW – has one of the lowest likelihoods of the alternatives tested.[18]

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 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 world cycle with factor loadings for most countries 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 with its 68 percent probability bands and the NBER recessions. Figure 3 shows the posterior inclusion probabilities for this cluster. The darkest areas indicate countries which are very likely to be included in this cluster and yellow indicates countries that are very likely not associated with the cluster. Countries in white are not included in our sample. Note, in particular, that cluster 1 does appear to demonstrate some regional/geographic properties. The cluster includes, with high probability, Japan and many of the countries in Europe. Other European countries – e.g., Iceland and Ireland – belong with more than 50 percent

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

probability. Brazil, Thailand, and Pakistan also belong with more than 50 percent probability. Not all 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 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 probability most of the countries in South America, with the exception of Brazil. Mexico, the Philippines, and a few African countries also belong to this cluster with high probability.

In contrast to a purely continental approach such as that used in KOW, our results suggest that a country such as Mexico is much more likely to have cycles similar to its shared-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 a country's investment share of GDP) is also important in determining the clusters. Similarities in countries' legal systems and 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. Again, the results are not directly comparable with those of KOW, but a number of qualitative similarities and differences highlight the effect of estimating the clusters. KOW find that, in general, the global factor explains a greater portion of the volatility in 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 output fluctuations of the 60 countries). Our results suggest a much larger role for the "regional" factor if a region is estimated by the countries' cyclical commonality. In fact, our cluster factors explain an average of 22.8 percent of the output fluctuations among the sample countries.

There are a few reasons this difference may not be surprising. First, KOW's sample differs from ours. In a recent paper, Hirato, Kose, and Otrok (2011) find that the importance of regional factors has increased over time. Because our sample includes more recent data than KOW, the regional factors have more explanatory power. Second, 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 the ability of their regional factor to explain output fluctuations. Third, imposing (rather than estimating) the regions may lead to the same misspecification discussed in the previous MC experiments above. When countries are included in a region with countries with which they do not actually share a common factor, the factor – and the associated loadings – may be biased.

Indeed, when the model is estimated with only output, 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 that 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 that groups them together, the regional factor explains 57 percent of their output variation. This increase in explanatory power is important, especially given that these countries account for a substantial share of the total output of the 60 countries in the sample.

6 Conclusions

Much 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 the third approach (that of KOW) estimates a factor model in which the correlation structure across countries is assumed to be determined by geographic proximity – that is, countries that 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 that can affect how countries are grouped. In MC experiments, we show that misspecifying the regions can affect the fit of the model. In the data, we find evidence that sharing a common geographic region is one component but not the *only* determinant of country groupings. These results, therefore, verify some of the underlying rationale behind KOW's selection of using a 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 that 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 majority of South American countries appear 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.

Priors for Estimation			
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*	$\kappa_0 = \frac{1}{M}$	$\forall n$
$\boldsymbol{\phi}$	$N(\mathbf{v}_0, \mathbf{V}_0)$	$\mathbf{v}_0 = \mathbf{0}_{p_F}, \mathbf{V}_0 = \frac{1}{2}\mathbf{I}_{p_F}$	$\forall i$
$\boldsymbol{\psi}$	$N(\mathbf{w}_0, \mathbf{W}_0)$	$\mathbf{w}_0 = \mathbf{0}_{p_\varepsilon}, \mathbf{W}_0 = \frac{1}{2}\mathbf{I}_{p_\varepsilon}$	$\forall n$
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. Notes: n denotes the series, where N is the total number of series; i indicates the cluster, where M is the total number of cluster factors; and the p's denote the maximum number of lags in the error and factor lag polynomials.

Degree of Cluster Misspecification									
	60%	40%	20%	6.7%	5%	3.4%	1.7%	*None*	*Endogenous*
Entropy	3372.2	3339.4	3302.7	3299.76	3295.30	3291.49	3289.98	3288.80	3287.45

Table 2: Monte Carlo Results. Notes: The table reports the median entropy for 1000 Monte Carlo replications with sample size of 50 periods. Each sample contains 60 series, 5 cluster factors, and 1 global factor. The column headings indicate the percent of the series in the exogenous clusters that is misallocated. 'None' indicates the exogenously clustered model with no misspecification. 1 misallocated series (of 60) equates to 1.66 percent misspecification, and so on. The last column shows the median entropy for the model estimated with clusters determined endogenously.

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.

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

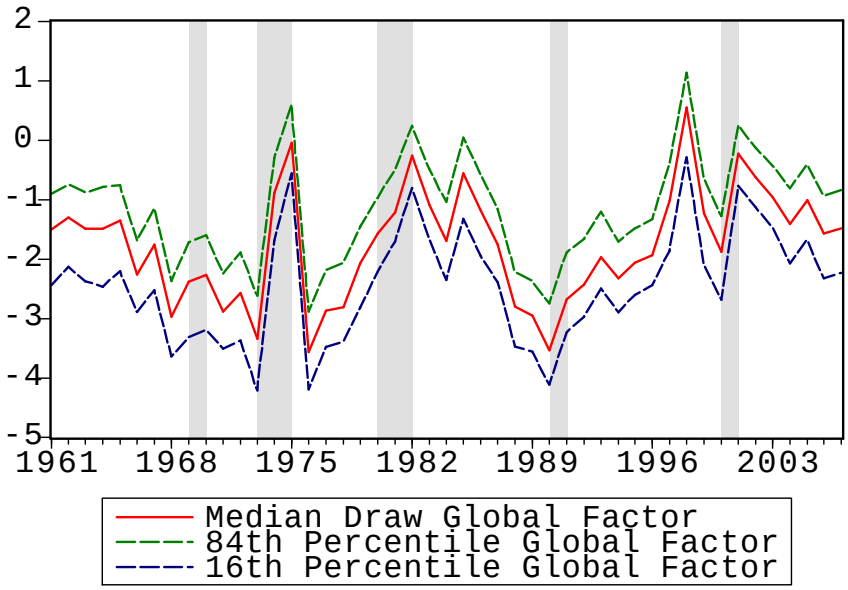
Table 4: Model Choice. Notes: The table shows the log marginal likelihood for model with various numbers of clusters estimated with the empirical data. The third column shows the difference in the log marginal likelihoods between the best model and each other model. The last column shows how much more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

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)

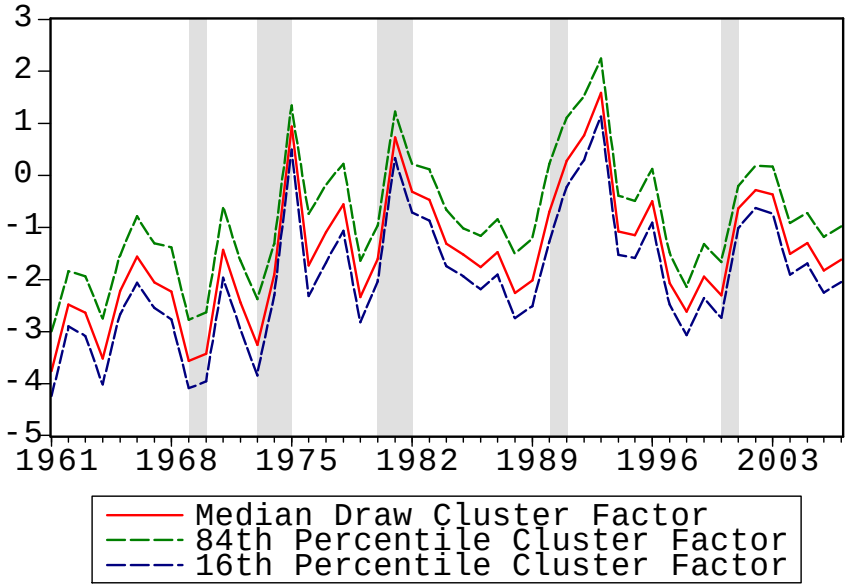
Table 5: Posterior means for each covariate in clusters 2 and 3. Notes: The first cluster (Cluster 1) covariate coeffients are normalized to zero. Values in bold indicate coefficients for which zero is not within the 68 percent coverage interval. The numbers in parentheses indicate the 16th and 84th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

Variance Decompositions: Posterior Means (%)							
	Global	Cluster	Idio.		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	Luxembourg	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. Notes: 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 are used to determine the cluster to which a country belongs, then the variance attributable to that specified cluster is calculated.



Global Factor. 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.



Cluster 1 Factor. 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.

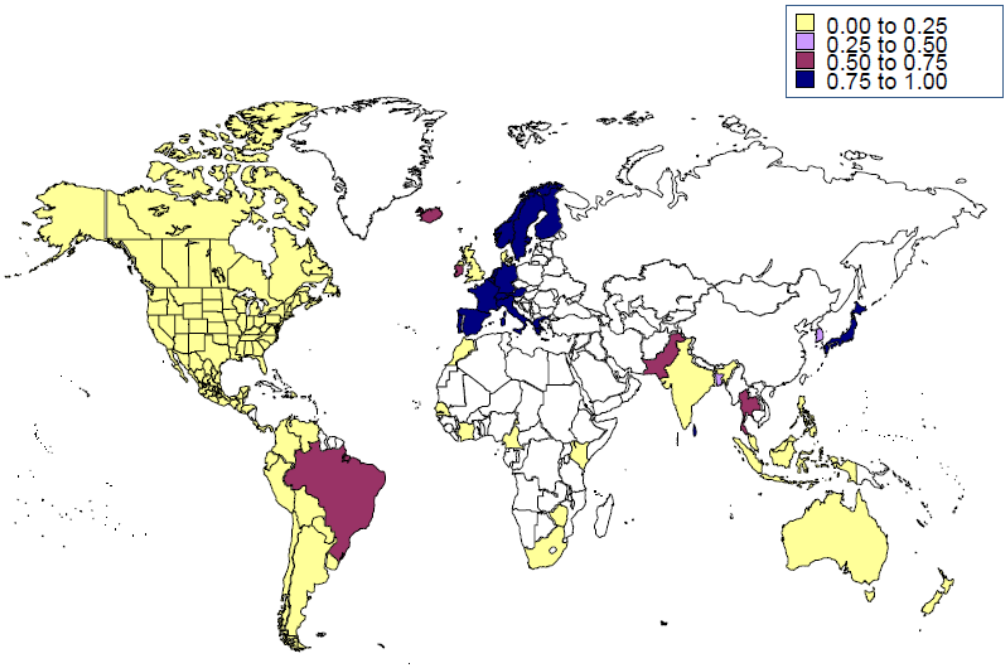
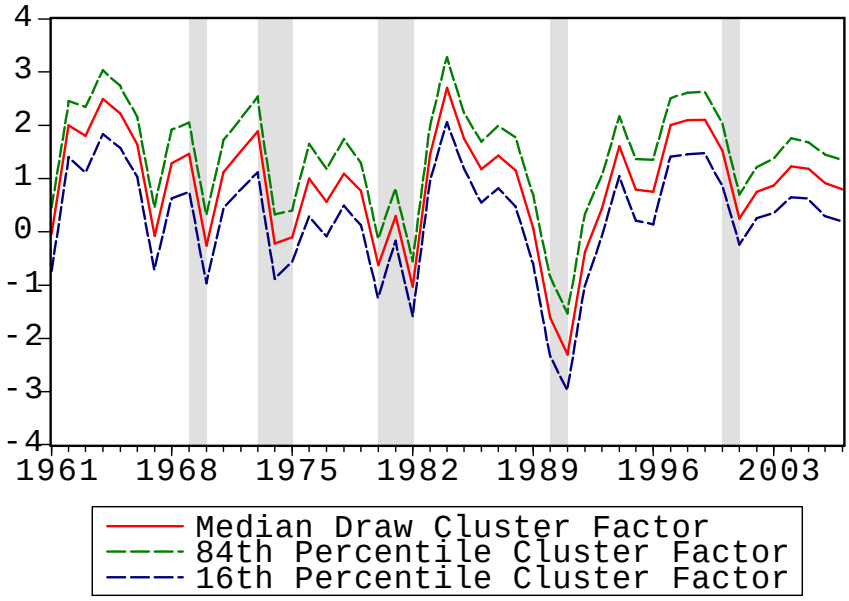


Figure 1: Cluster 1 Composition. The map shows the posterior probabilities of countries included in Cluster 1. Countries in white are omitted from the sample.



Cluster 2 Factor. 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.

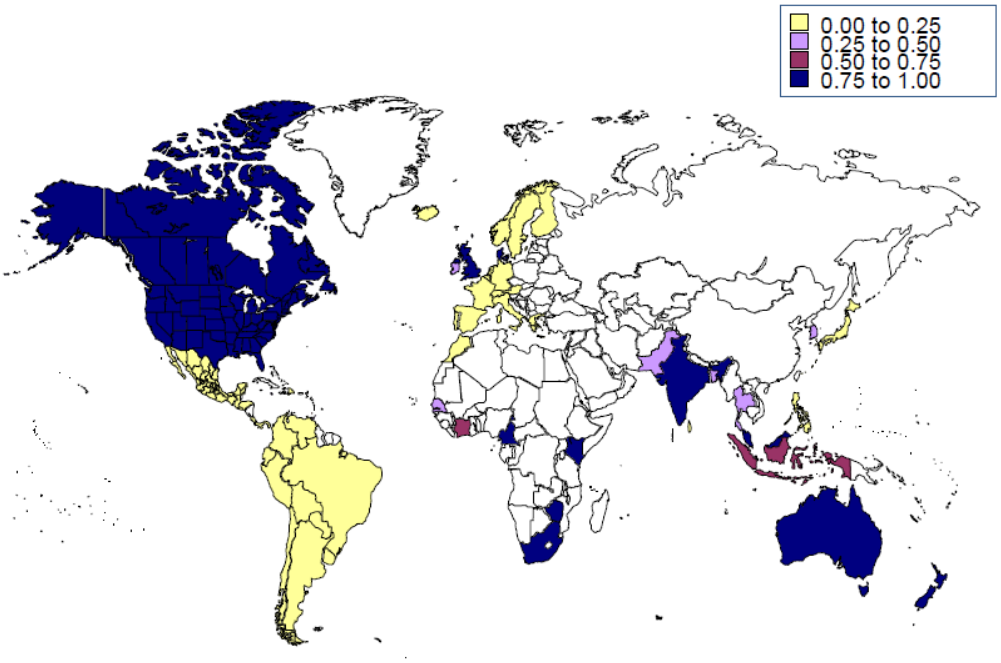
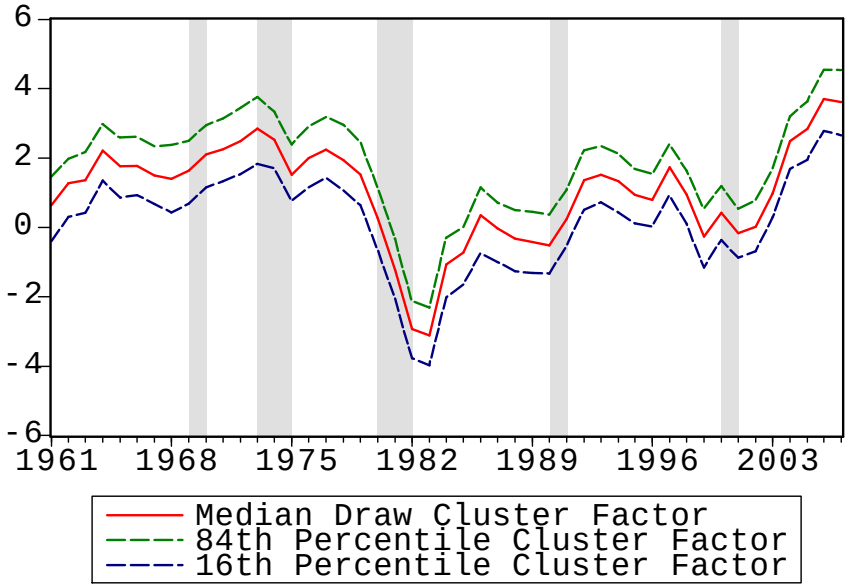


Figure 2: Cluster 2 Composition. The map shows the posterior probabilities of countries included in Cluster 2. Countries in white are omitted from the sample.



Cluster 3 Factor. 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.

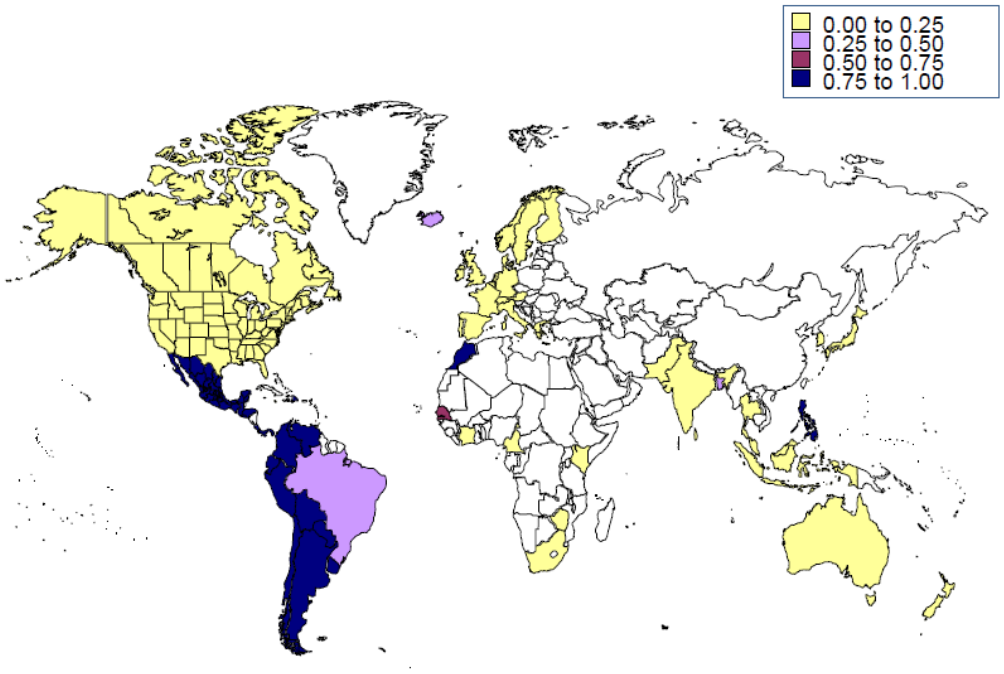


Figure 3: Cluster 3 Composition. The map shows the posterior probabilities of countries included in Cluster 3. Countries in white are omitted from the sample.

References

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

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

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

[4] 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

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

[6] Carter, C.K. and Kohn, R. "On Gibbs Sampling for State Space Models." *Biometrika,* August 1994, 81(3), pp. 541-53.

[7] 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.

[8] Casella, George and George, Edward I. "Explaining the Gibbs Sampler." *American Statistician,* August 1992, 46(3), pp. 167-74.

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

[10] Chib, Siddhartha. "Marginal Likelihood from the Gibbs Output." *Journal of the American Statistical Association*, December 1995, 90(432), pp. 1313-21.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

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

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

[17] Djankov; La Porta; Lopez-de-Silanes; and Shleifer. "Courts." *Quarterly Journal of Economics*, May, 2003

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

[19] 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.

[20] 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.*

[21] Frühwirth-Schnatter, Sylvia and Kaufmann, Sylvia. "Model-Based Clustering of Multiple Times Series." *Journal of Business and Economic Statistics*, January 2008, 26(1), pp. 78-89.

[22] Frühwirth-Schnatter, Sylvia and Lopes, Hedibert Freitas. "Parsimonious Bayesian Factor Analysis when the Number of Factors Is Unknown." Unpublished, 2009.

[23] 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.

[24] 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.

[25] Hallin M. and R. Liska. "Dynamic factors in the Presence of Block Structure," EUI Working Papers, 2008, ECO 2008/22.

[26] Hamilton, James D. and Owyang, Michael T. "The Propagation of Regional Recessions." WP (2009), forthcoming, *Review of Economics and Statistics.*

[27] Harati, Hideaki; KOse, M. Ayhan; and Otrok, Christopher. "Regionalization vs. Globalization," manuscript.

[28] 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.

[29] Holmes, Chris C. and Held, Leonhard. "Bayesian Auxiliary Variable Models for Binary and Multinomial Regression." *Bayesian Analysis,* 2006, 1(1), pp. 145-168.

[30] 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.

[31] 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.

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

[33] 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.

[34] 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.

[35] La Porta; Lopez-de-Silanes; Shleifer; and Vishny. "The Quality of Government." *Journal of Law, Economics and Organization*, 1999.

[36] McKinnon, Ronald I. "Currency Substitution and Instability in the World Dollar Standard." *American Economic Review*, June 1982, 72(3), pp. 320-33.

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

[38] 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.

[39] 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

[40] 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.

[41] 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.

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

[43] 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.

[44] 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