Is the US an OCA?

Patrick M. Crowley
College of Business,
Texas A&M University,
Corpus Christi,
Texas, USA

November 2001

ALL COMMENTS WELCOME!

Paper to be presented at the Southern Economic meetings in Tampa, Florida, USA, November 16th-18th, 2001.

Affiliations:

Texas A&M University, College of Business, Corpus Christi, Texas, USA.
Fellow, Athenian Policy Forum.

Contact details:

Patrick Crowley,
College of Business,
Texas A&M University,
Corpus Christi,
Texas, 78412
USA.
Tel: (361) 825-6011
Fax: (361) 825-5609
email: pcrowley@cob.tamucc.edu

JEL Classification:

F15, F31

Acknowledgements

I would like to thank my research assistant Josh Osborne at Texas A&M for excellent assistance over the past 6 months and Dean Abdelsamad for funding my research last year. I would also like to thank Chris Fraley and Adrian Raftery of the Department of Statistics at University of Washington for making the MCLUST model based software available.

1. Introduction

Economic theory, following the pioneering work of Mundell (1961) and others, suggests that regions that are subject to similar economic cycles and shocks could be subject to the same monetary policy and hence use the same currency - hence the notion of optimal currency areas (OCAs). Operationalizing the concept of an OCA has been a challenge to economists, and there have been several interpretations as to what stylized empirical observations should determine what does and what does not constitute an OCA.

The OCA approach was particularly important in the case of the EU, which has embarked on the ambitious project of replacing national currencies with a single EU currency, the euro. In the EU context, much ink has been spilled over both optimal currency areas (OCA) and the economic convergence criteria that were used to determine membership of the single currency project (see Crowley (1996), Buiter, Corsetti and Roubini (1993), Fratianni, von Hagen, and Waller (1992), Eichengreen (1993), to name but a few), but little work has been done on the applicability of these criteria to other regional integration agreements.

This paper seeks to evaluate whether the US in its current form forms an optimal currency area, in the sense that i) synchronicity of business cycles is achieved; ii) states have a similar experience with movements in inflation and unemployment; and iii) that levels of labor mobility and labor market conditions are similar. The paper is divided into six sections. Section 2 evaluates the theoretical and empirical literature on optimal currency areas and business cycle synchronicity, while section 3 reviews the methodology used in this paper. Section 4 outlines the data and data preparation. Section 5 presents the empirical results of the analysis of business cycle correlations and other associated variables using model-based cluster analysis. Section 6 concludes.

2. OCAs and Business Cycles

a. The Optimal Currency Area Literature

Two seminal papers on optimal currency areas (Mundell (1961) and McKinnon (1963)) outlined the conditions under which several administrative jurisdictions might be suitable to be subject to the same monetary policy. Further refinements of this approach were subsequently made by Kenen (1969). Bayoumi (1994) also offered a formal model of optimal currency areas (OCAs) with microeconomic foundations to underscore Mundell's original thesis. The conditions for an OCA are that members of the currency union should, for the most part, experience symmetric shocks and that economic cycles should be synchronous. If countries experience asymmetric shocks or have asynchronous business cycles then the costs of being subject to a single monetary policy may be significant, and may outweigh the costs. To offset asymmetric shocks or asymmetric business cycles, then certain currency area characteristics may ameliorate costs, notably i) a significant degree of labour mobility, ii) fiscal transfers through a "federal" level of government and iii) flexible wages and prices (see Melitz and Zumer (1998)). Part of the reason why the OCA literature has been such a focus of interest in the context of the EU has been due to the fact that "euroland" cannot be characterised as possessing these characteristics to the same degree that the United States does, and that participation in European Economic and Monetary Union (EMU) was determined largely by satisfaction of economic criteria.

The empirical time-series literature on OCAs can be divided into three strands - a strand that uses basic regional data (from a sub-national level) to evaluate whether countries use exchange rates to offset shocks, with the implication that similar exchange rate volatilities would imply similar shock magnitudes, while at the same time evaluating whether participants possess the three offsetting characterstics (see De Grauwe and Vanhaverbeke (1993)); a strand that uses structural vector autoregression (SVAR) time series methodology (following Blanchard and Quah (1989)) to identify demand and supply shocks (see for example, Bayoumi and Eichengreen (1994a) for the EU and North America and Lalonde and St-Amant (1993) and DeSerres and Lalonde (1994) for Canada) and then look at the correlation of these shocks across countries or regions. Lastly, another strand of the literature evaluates the synchronicity of business cycles across prospective currency union members (Baxter and Stockman (1989) and Artis and Zhang (1997a)).

The first strand of OCA empirical research has been criticised for being largely descriptive, while the second (SVAR) methodology has been criticised (see Buiter (1998)) for being arbitrary in terms of the restrictions that are required for identification of monetary and real shocks (usually the assumption that shocks that are neutral in the long run are monetary shocks). The third strand of research also responds to another criticism of the VAR methodology: that a shock approach ignores long run business cycle synchronicity - the synchronicity approach compares cyclical components in GDP and then uses correlations in business cycles to draw out implications about suitability as constituents of an OCA. The obvious drawback here is that this approach completely ignores the incidence of temporary shocks and does not consider the ability of exchange rates to also compensate for shocks.

A recent development in the OCA literature has been recognition that ex-ante evaluations of which countries constitute an OCA might ignore the Lucas critique, in that new members of an OCA might a) modify policy to be better suited to an OCA (see Tavlas, G. (1993)) or b) be more suited to being in an OCA *ex-post* (Frankel and Rose (1997)). The latter approach takes into consideration factors which usually do not appear in the *ex-ante* OCA approach, such as trade intensity, real interest rate cycles, and fiscal policy coordination[1].

[1] Neumeyer (1998) also considers the notion that political shocks could be incorporated as another variable contributing to the factors which might suggest an optimal currency area.

An excellent survey of recent developments in the optimal currency area literature can be found in LaFrance and St-Amant (1999).

b. Synchronicity of Business Cycles

Following the work of Gerlach (1988) and Baxter and Stockman (1989) on business cycle correlations, there has been considerable research devoted to the propagation of business cycles, and the existence of a world business cycle in the pre- and post- Bretton Woods periods. Recent research on business cycles has focussed on the effects of trade in propogating business cycles (see Imbs (1999)) and on new measures of co-movement (see Croux, Forni and Reichlin (1999)) of output data for different regions or countries.

Artis and Zhang (1997a) explored the idea of group-specific business cycles after the inception of the ERM of the EMS in 1979, positing a distinctly European business cycle[2]. In this study, cyclical components of industrial production were obtained using several de-trending methods[3], and then the cross-correlations of the cyclical components of these series with the US series and the German series were calculated. A European business cycle was confirmed, but the cycle was confined to members of the ERM of the EMS, as might have been expected. The results were shown to be robust to the detrending method employed.

Here we employ the same similar methodology, with two differences. First, in the European context, Artis and Zhang (1997a) justified using the cyclical component of the German series as a basis for evaluating whether a European business cycle existed, predicated on other research which clearly

---

[2] Further research by Artis, Krolzig and Toro (1999) has analysed the phasing of the European business cycle.

[3] The methods used were phase-average trend (PAT) detrending (Nilsson, 1987), a linear trend and a Hodrick-Prescott filter.

showed Germany to be the largest and most influential economy in the EU, and the Bundesbank to be a "leader" in terms of the setting of monetary policy in the ERM of the EMS. In the US context there is already a single monetary authority, so the U.S. national aggregates are used as the appropriate "target" variables for the purposes of calculating cross-correlations for individual US states. Second, as research on Canada and NAFTA (see Crowley (2000 and 2001)) showed that some lagged effects on regional GDP took place, also a lagged business cycle correlation is calculated.

The analysis was undertaken by estimating cyclical movements using a Hodrick-Prescott filter. A Hodrick-Prescott filter is based on minimizing the following expression with respect to g_t:

$$\min\left\{\sum_{t=1}^{N}(y_t - g_t)^2 + l\sum_{t=2}^{N-1}[(g_{t+1} - g_t) - (g_t - g_{t-1})]^2\right\} \tag{1}$$

where y_t is the raw data, g_t is the growth component, and so (y_t - g_t) is the cyclical component. 8 is a dampening parameter whose value is extremely sensitive to the data being used. Following recent research by Pedersen (2000), which calculated optimal values of 8 using the estimated spectral shape of real US GDP in both monthly, quarterly, and annual series, a value of 8=40 was used for annual data and a value of 8=800 was used for monthly data. Figure 1 below shows correlations for the cyclical component of GDP.

Figure 1

Contemporaneous and Lagged Business Cycle Cross-Correlations

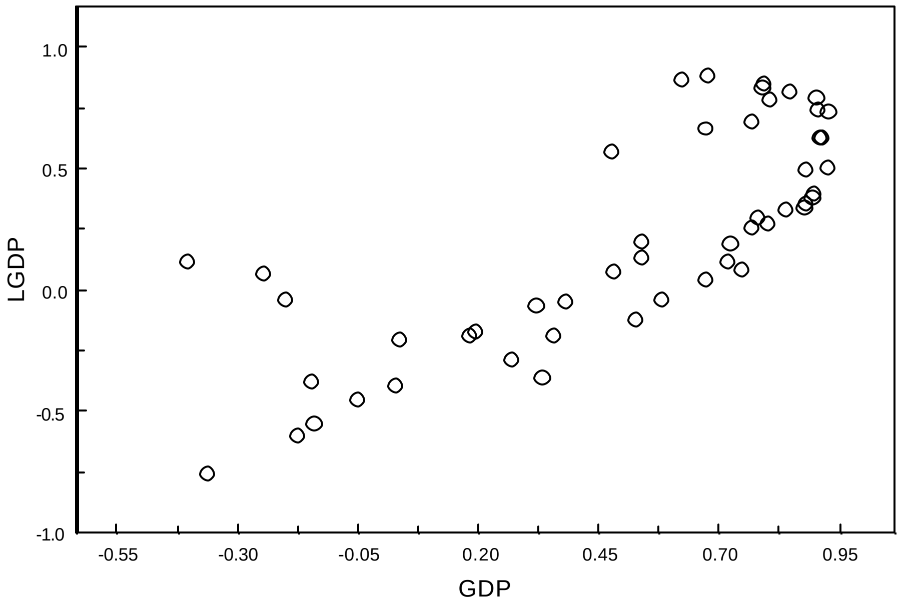
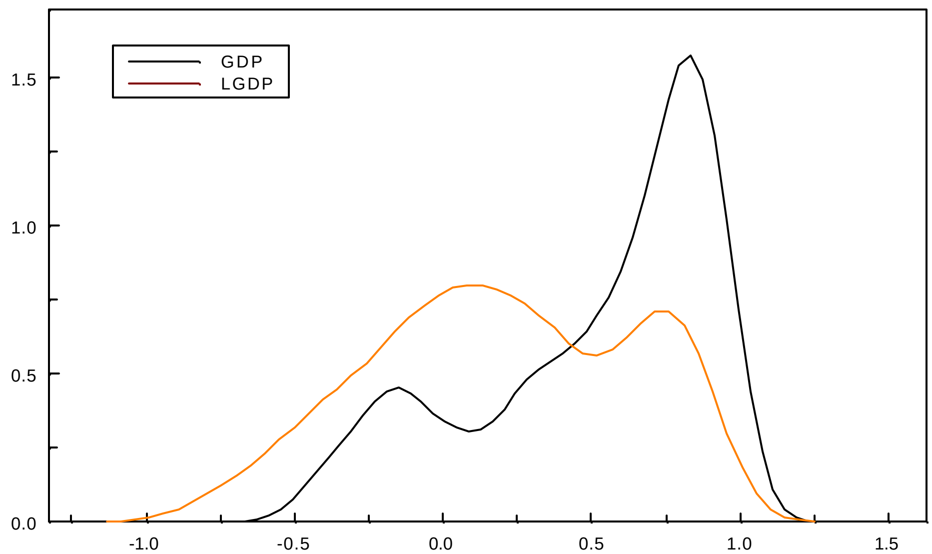


Figure 1 shows that most states have a high contemporaneous cross-correlation with US GDP, whereas, with a few exceptions, most states do not have such a high one-year lag cross-correlation. Using kernel density estimation, this can easily be seen in figure 2 below:

Figure 2

Kernel Density Estimation for Business Correlations



It is clear from the estimated densities, that many states have high contemporaneous cross-correlations with the US where as most states have a positive lagged correlation coefficient, but given the shapes above, the majority will have a contemporaneous correlation coefficient that is greater than the lagged correlation coefficient, but the tail of the lagged distribution does exhibit higher density than for the contemporaneous distribution. For completeness, annex B tabulates the correlation coefficients used in the analysis by state.

3. Methodology

a. Cluster analysis

Cluster analysis was first applied by Fisher (1936) to classifications of irises (found by Anderson (1935)) indigenous to the Gaspé peninsula in Québec. In economics cluster analysis has been applied to EU data by several authors, notably Jacquemin and Sapir (1995) and Artis and Zhang (1997b and 1998a and b), with interesting results. The cluster analysis done on the EU has largely corroborated the evidence on suitability for membership of EMU gained from the aforementioned empirical methods used in the OCA literature. Artis and Zhang (1998a) found three EU groups of Member States using this approach based on synchronisation in business cycle, real interest rate cycle, volatility in real exchange rate, openness to trade, convergence of inflation and labour market flexibility. The methodology is also started to infiltrate into the economics profession in North America, with Galbraith and Jiaquing (1999), Honohan (2000) and Crone (1999) using the technique - also Maharaj and Inder (1999) is another recent application using cluster analysis to forecast time series in economics. In other disciplines, cluster analysis is frequently used - applications range from astrophysics (Mukerjee, Feigelson, Babu et al (1998)) to microbiology (van Ooyen (2001)).

Cluster analysis aims to determine the intrinsic structure of data when no information other than the observed values is available - the data is to be partitioned into meaningful subgroups. This approach should be put in contrast with discriminant analysis, in which known groupings of some observations are used to categorize others and infer the structure of the data as a whole. Clustering methods range from those that are largely heuristic to more formal procedures based on statistical models, and they are hierarchical or based on allocating observations among tentative clusters (such as k-means clustering).

Hierarchical methods proceed by stages, partitioning or combining the data at each stage. Hierarchical methods fall into two categories: “agglomerative” and “divisive” - with agglomerative denoting the merging of clusters at each stage and divisive denoting the splitting of clusters at each stage - in most cases agglomerative and divisive methods give similar clusterings. At each stage some criterion is optimized used to determine which clusters should be combined or split - most methods use single link (nearest neighbor), complete link (farthest neighbor) or sum of squares. In model methods, however, usually a maximum likelihood based on specific distributional assumptions is used to merge or divide groups. Useful references for these heuristic clustering methods are Anderberg (1993), Kaufman and Rousseeuw (1990) and Hartigan (1975).

Unfortunately, although these clustering methods are appealing, none of them addresses the issue of how many clusters there should be. Various strategies have been put forward to choose the number of clusters, but up until recently none of these methods has been satisfactory from a computational point of view, or from a methodological point of view (see Bock (1996) for a survey of this issue and related research). The alternative that has been presented by Fraley and Raftery (1998a and b) is computationally relatively straightforward, and is also intuitively appealing, so this methodology is adopted here.

b. Model-based cluster analysis

In probability based clustering, each observation is assumed to be generated by a mixture of underlying probability distributions where each component in the mixture represents a different cluster. Given a set of observations $\mathbf{x} = (\mathbf{x}_1,\dots,\mathbf{x}_n)$, then the density of an observation $\mathbf{x}_i$ from the kth component in a total number of G components, is $f_k(\mathbf{x}_i * 2_k)$, where 2_k are the parameters. In most cases, $f_k(\mathbf{x}_i * 2_k)$ is assumed to be multivariate normal (Gaussian), so in this instance the parameters 2_k consist of a mean vector $:_k$ and a covariance matrix G_k. The clusters will then be ellipsoidal, with center at $:_k$, and the covariance matrix will determine the other characteristics.

The mixture likelihood approach then maximizes the criterion:

$$\ell_M(\theta_1,\dots,\theta_G;\tau_1,\dots,\tau_G \mid x) = \prod_{i=1}^{n} \sum_{k=1}^{G} \tau_k f_k(x_i \mid \theta_k) \qquad (2)$$

where J_k is the probability that an observation belongs to the kth component.

Banfield and Raftery (1993) developed a model-based framework for clustering by expressing the covariance matrix in terms of its eigenvalue decomposition, which is of the form

$$\sum_k = \lambda_k D_k A_k D_k^T \tag{3}$$

where D_k is the orthogonal matrix of eigenvectors, A_k is a diagonal matrix where the elements of the diagonals are proportional to the eigenvalues of G_k, and 8_k is a scalar. This leads to a geometric interpretation of the ellipsoidal clusters - D_k determines the orientation, A_k determines the shape of the density contours and 8_k specifies the volume. These characteristics can then be allowed to vary between clusters, or constrained to be the same for all clusters. This approach actually subsumes many previous approaches at model-based clustering - more details can be located in Fraley and Raftery (1998a). The range of models used here is limited, given the limitations of the SPLUS software and the library MCLUST which was used for estimation - a more extensive set of models within the same framework can be found in Celeux and Govaerts (1995).

In the approach taken here, the parameterizations of the covariance matrix are detailed in table 1 below:

Table 1

Parameterizations of the Covariance Matrix

Model	**ID**	**Distribution**	**Volume**	**Shape**	**Orientation**
$8I$	EI	Spherical	Equal	Equal	NA
8_kI	VI	Spherical	Variable	Equal	NA
$8DAD^T$	EEE	Ellipsoidal	Equal	Equal	Equal
$8_kD_kA_kD_k^T$	VVV	Ellipsoidal	Variable	Variable	Variable
$8D_kAD_k^T$	EEV	Ellipsoidal	Equal	Equal	Variable
$8_kD_kAD_k^T$	VEV	Ellipsoidal	Variable	Equal	Variable

Source: Banfield and Raftery (1993)

Given the different model parameterizations above, agglomerative hierarchical clustering can be used by merging clusters so as to maximize the resulting likelihood as specified in equation (2) above.

c. Clustering algorithms

The algorithm used for maximizing the likelihood function here is the EM (Expectation-Maximization) algorithm (see McLachlan and Krishnan (1997)). EM iterates between an "E" step, which computes a matriz z such that z_{ik} is an estimate of the conditional probability that observation i belongs to group k given the current parameter estimates, and an "M" step, which computes maximum likelihood parameter estimates given z. In the limit, under certain conditions the parameters usually converge to the maximum likelihood values for the Gaussian mixture model and the sums of the columns of z converge to n times the mixing proportions J_k, where n is the number of observations.

The EM algorithm is not without its problems though. Banfield and Raftery (1998a) detail several problems notably i) a slow rate of convergence, ii) the number of conditional probabilities associated with each observation equals the number of components in the mixture, so that the EM algorithm may not be suitable for large datasets and iii) when the covariance matrix becomes singular or nearly singular (otherwise known as “ill-conditioned”) the EM algorithm breaks down. The latter problem was an issue in this study - usually relates to clusters which only contain a few observations or if the observations contained are co-linear, and in this study the former is the suspected problem.

d. Model selection

The mixture model approach allows the use of approximate Bayes factors to compare models (see Kass and Raftery (1995)). The Bayes factor is the posterior odds for one model against the other assuming neither is favored a priori. With the EM algorithm twice the log Bayes factor is used to determine the number of clusters in hierarchical clustering based on the mixture likelihood.- this measure is also known as the Bayesian Information Criterion (BIC) and is specified as:

$$2\log p(x|M) + \mathrm{const} \approx 2\ell_M \log(n) - m_M \log(n) \equiv \mathrm{BIC} \qquad (4)$$

where $p(x|M)$ is the likelihood of the data for the model M, $l_M(x|\Theta)$ is the maximized mixture log likelihood for the model and m_M is the number of independent parameters to be estimated in the model. The larger the value of the BIC, the stronger the evidence for the model[4].

A standard convention for calibrating BIC differences is that differences of less than 2 correspond to weak evidence, differences between 2 and 6 to positive evidence, differences between 6 and 10 to strong evidence, and differences greater than 10 to very strong evidence.

[4] This is because the number of clusters is not considered independent for calculating the BIC, and hence if each model is equally likely, the posterior probability p(x|M) should be higher, and hence the BIC should be higher.

e. Clustering strategy

The general strategy adopted here is similar to that of Fraley and Raftery (1998a) and is detailed for SPLUS library MCLUST in Fraley and Raftery (1998b). The steps of strategy are as follows:\

i) determine a maximum number of clusters to consider, and a set of candidate parameterizations of the model to use.

ii) use agglomerative hierarchical clustering for the unconstrained Gaussian model, to obtain classifications for up to M groups.

iii) do EM for each parameterization and each number of clusters, starting with the classification from hierarchical clustering.

iv) compute the BIC for the one cluster model for each parameterization and for the mixture likelihood with optimal parameters from EM for other clusters.

v) plot the BIC - this should hopefully indicate a local maximum and a specific model.

This strategy was followed for the research presented here.

4 Data and data sources

To use cluster analysis for the US following the optimal currency area theory, data is needed that corroborates the degree of synchronicity in business cycles plus the degree in flexibility in labor markets. State data for the US is available on a limited basis, so variables are selected or constructed to attempt to best characterise the flavor of what the optimal currency area literature suggests should be important in subjecting regions to a common monetary policy. Annex C documents the data sources.

Here we use both the contemporaneous and lagged cyclical business cycle cross-correlations from section 2. In addition to these two variables, we include:

i) unemployment rate correlations - this is used as another business cycle correlation variable, although it could also proxy both the degree of labor market flexibility compared to the US, plus it can also be thought of as a measure of labor migration inertia;

ii) inflation rate correlations - this is the CPI by state or province and is used to capture commonality in consumer price inflation experience (these are obtained from Leonard and Walder (1999));

iii) labor absorption/leakage measure - this is population minus births plus deaths as a measure of inward/outward labor flow. This is summed over the total period, and then the absolute value of the sum is correlated against the US value;

iv) average hourly wages - these are detrended using a Hodrick-Prescott filter and then correlated against the US cyclical component.

The data used to construct the variables above are detailed in annex C. The above gives us 6 pieces of economic data to use for cluster analysis for each state, giving a data set of 300 observations.

Each annual series was correlated against its U.S. national counterpart, and then normalized as is standard in cluster analysis. Annex D gives details of the periods over which cor relations were calculated. Figure 3 below shows a scatterplot for the migration measure and average wage correlations for the data.

Figure 3

Scatterplot of Average Wage and Migration Correlations

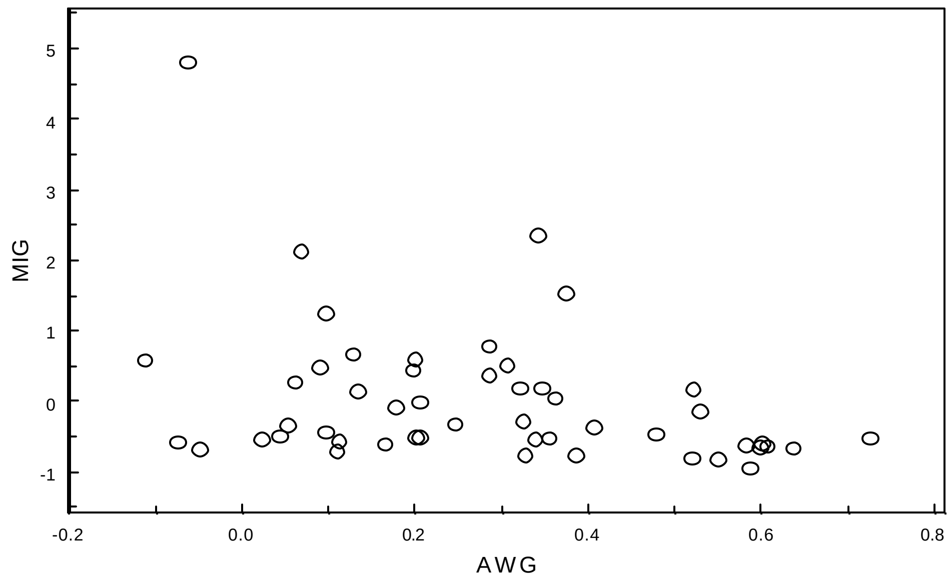


Figure 3 shows that although many states have differing correlations with the US for average wage rates, the less correlated they are with the US, the more likely they are to have a greater degree of in- or out-migration.

One way of displaying the data used here is to use a method that in statistics is known as "Chernoff's faces" (see Chernoff (1973)). Here faces are used to describe the correlations, where the feature parameters are as follows: the larger the area of the face, the higher the correlation with US GDP, the more rotund the shape of the face, the larger the correlation with lagged GDP, the longer the length of the nose, the higher the correlation with US unemployment rates, the more central the location of the mouth, the higher the correlation with US inflation rates, the more curved the smile the larger the correlation of average wages and the wider the mouth the higher the level of labor in- or out-migration compared with the US.

Figure 4

Chernoff's faces for US data

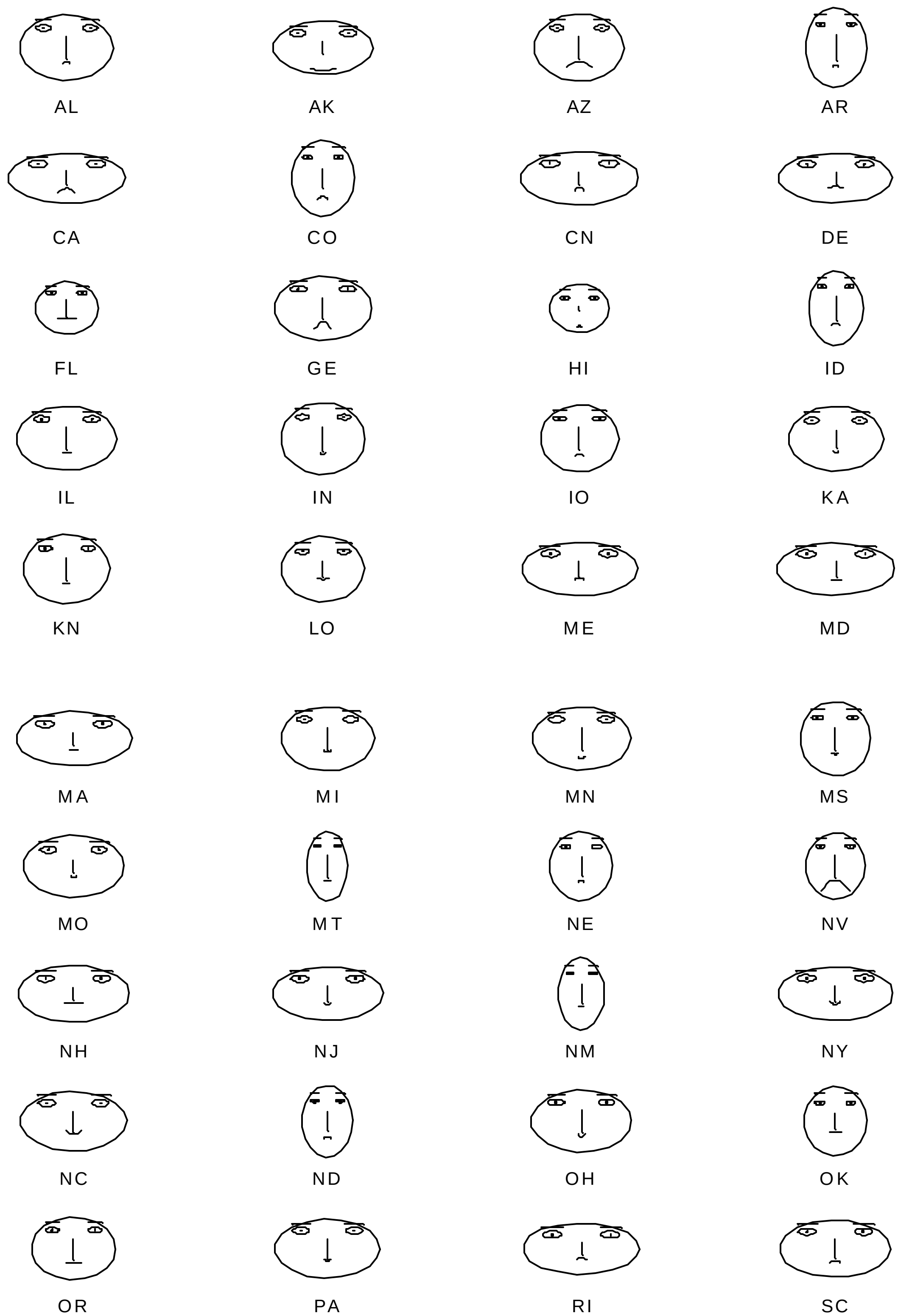
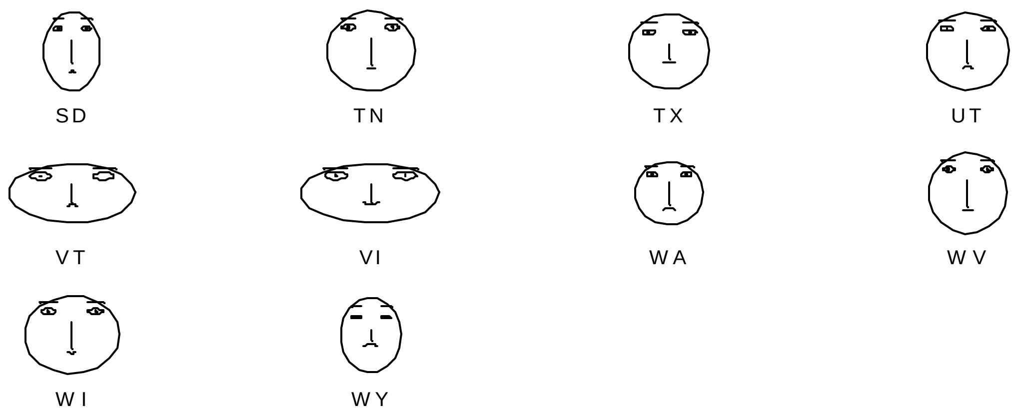


Figure 4 (continued)



The faces clearly show a wide variation between U.S. states, with certain states like Hawaii (HI) and Wyoming (WI) appearing to very different from the rest of the country. Interestingly as well, nearly all the faces have small mouths - this reflects the fact that most of the states do not have significantly different degrees of migration to the US, and so although the correlations with the US average wage rate are mostly greater than zero (- in other words the faces should largely be smiling) the lack of difference with the US on labor mobility hides this fact (- the mouths are too small to detect a smile!). Even at this stage, given our exploratory data analysis, it appears apparent that certain states have had a rather different experience from the rest of the US over the period in question, so might be expected to fall into separate categories (clusters).

For cluster analysis, correlation coefficients were not normalized as their range is no larger than (-1,+1). The exception here was the migration variable (which is not a correlation coefficient) - in this instance it made sense to standardize this variable. Annex D tabulates the means and standard deviations for correlations for each variable.

5 Empirical Results

As section 4 e) detailed, the cluster analysis methodology adopted here. In all cases the EM algorithm was initialized using hierarchical clustering using the unconstrained model (EI) detailed in table 1 above. The maximum number of clusters was chosen to be M=10. From this point BIC values were calculated from an initial parameterization for all other possible models presented in table 1. Table 2 gives the BIC for each of the candidate models for each of the 6 cluster groups specified. Some BIC estimates were not available, as the covariance matrix associated with one or more of the mixture components is ill-conditioned, so that the log likelihood and hence the BIC cannot be computed.

Figure 5 below then graphs the BIC plots by numbers of clusters.

Table 2

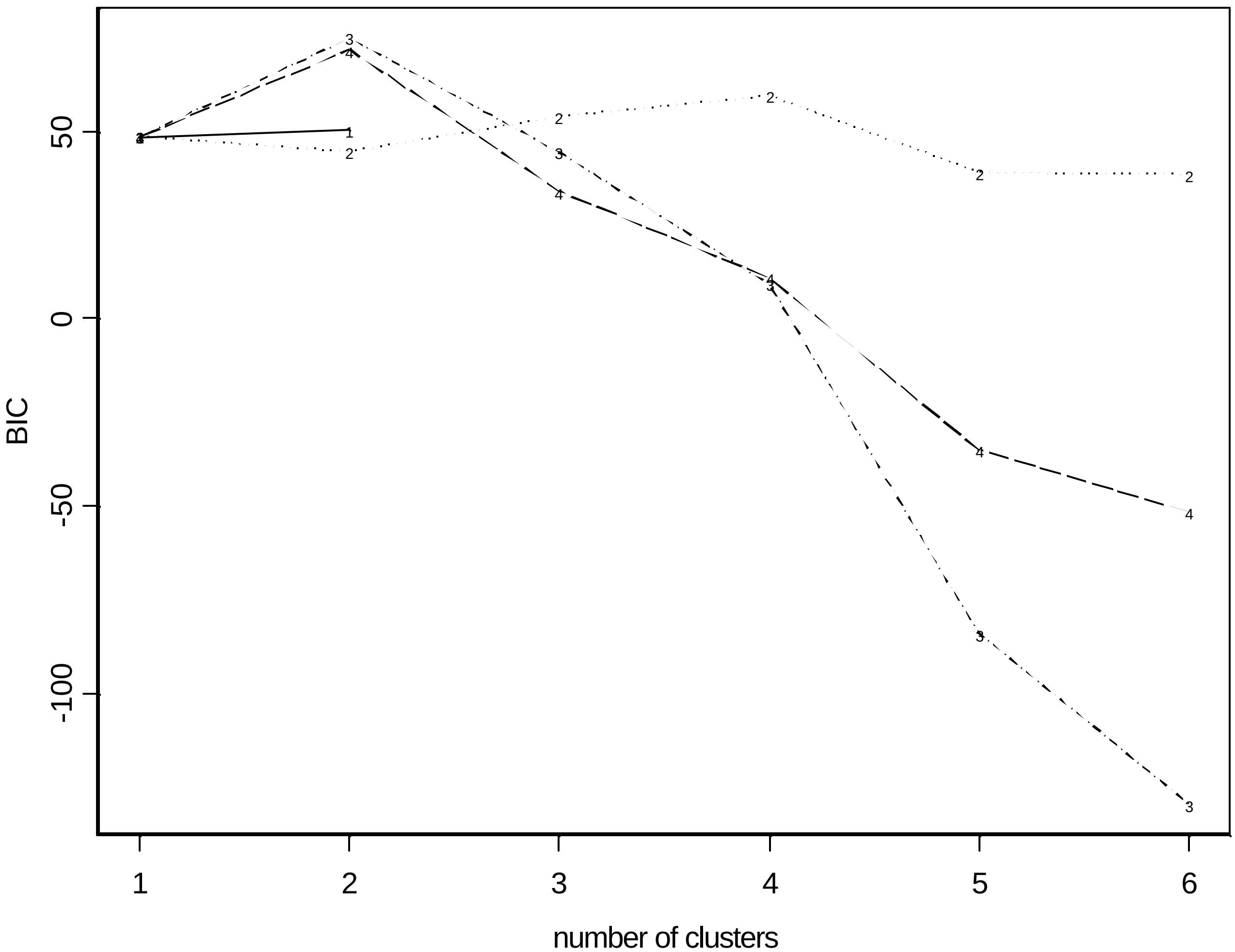
BIC Values by Number of Clusters using the EM algorithm

ID	**1**	**2**	**3**	**4**	**5**	**6**
EI	-445.37	-319.00	-261.97	-254.36	-247.95	-266.63
VI	-445.37	-280.97	-205.75	-189.61	-162.87	-153.54
EEE	48.53	44.57	53.91	59.46	39.09	38.43
VVV	48.53	50.40	NA	NA	NA	NA
EEV	48.53	74.82	44.38	9.22	-83.96	-129.45
VEV	48.53	71.66	33.90	10.81	-34.99	-51.47

NA = ill-conditioned matrix   Initialized using the EI model

Figure 5

BIC by Numbers of Clusters



Key: 1=VVV, 2=EEE, 3=EEV, 4=VEV

Figure 5 shows that the best BIC value is obtained for 2 clusters under EEV (ellipsoidal, constant shape, equal volume), with the second highest value for VEV (ellipsoidal, variable volume and orientation, but equal shape), with little to choose between the two different parameterizations. The difference in the BIC is 3.16, so that the evidence for EEV is positive, but not overwhelming. We therefore consider both parameterizations for the cluster results. The third highest BIC value was obtained for model VVV (ellipsoidal, variable shape, volume and orientation) again with 2 clusters.

Now that the optimum number of clusters has been determined, model based clustering can be implemented with 2 clusters under both the the EEV and VEV models using the EM algorithm. The results of the clustering are presented in table 3 below:

Table 3a

Cluster Membership using EEV model and EM algorithm

Cluster	States
1	AK, FL, HI, MT, NV, NM, ND, SD, WA
2	Rest of the states

Uncertainty for EEV (quantiles):

0%	25%	50%	75%	100%
0	0	0	1.8E-11	0.298

Table 3b

Cluster Membership using VEV model and EM algorithm

Cluster	States
1	AL, AK, AZ, AR, CA, CO, DE, FL, GE, HI, ID, LO, MO, MT, NE, NV, NH, NM, NY, NC, ND, OK, OR, SD, TX, UT, WI, WA WV, WY
2	CN, IL, IN, IO, KA, KN, ME, MD, MA, MI, MN, MS, NJ, OH, PA, RI, SC, TN, VT, WI

Uncertainty for VEV (quantiles):

0%	25%	50%	75%	100%
0	0	3.44E-11	0.0072	0.151

Coding given in Annex A

The uncertainty in the classification associated with the EM algorithm gives a measure of the quality of the classification. In this case, using examples given in Fraley and Raftery (1998b), the majority of observations are well classified. Uncertain classifications result when groups intersect, giving overlapping regions and therefore some uncertainty as to which cluster to allocate to. Looking at the uncertainties by state, it is clear that the VEV gives more uncertain points, but less uncertainty on average for each uncertain point, where as for the EEV model, there are less uncertain points, but more uncertainty on average for each uncertain point.

The clustering shown in tables 3a) and 3b) is clearly very different, with the EEV model (table 3a)) having a much easier interpretation than the VEV model (table 3b)). The numbering of the clusters indicates the order in which each cluster formed according to the agglomeration procedure.

Given that the EEV model is more likely, given the BIC results, then table 3a) clearly implies several empirical facts, given the data and time periods used in the analysis:

i) the U.S. largely constitutes an OCA as only 2 clusters were found using this methodology;

ii) peripheral states or remote states may not be asynchronous with the majority of states and hence may form a separate cluster;

The implications from this analysis are that the US does not form an optimal currency area. Of course the argument still remains that the US is a fiscal union, so that if federal transfers were included as offsetting items then these elements would render the US an OCA. Given the results above, and previous work on Canada, it implies that some other "glue" binds these two countries together other than economic cycles (e.g. fiscal transfers, history or political motives).

6 Conclusions

This paper asked the question is the US an OCA? On the basis of the data used here the answer is plainly no. The research used a model-based cluster analysis approach employing a new strategy using Bayesian techniques to choose the optimal number of clusters and the model to be used for the clustering. Further, the clustering algorithm used, the EM algorithm, which permits allocation of observations based on a maximum likelihood procedure. On the basis of the model-based clustering techniques used though, the results confirmed that the US consists of two different clusters. Membership between the two most likely candidates was not consistent though. The model given the highest likelihood indicated that states on the periphery or in more isolated parts of the US tended to form a cluster by themselves. Clearly other factors such as historical, political, or fiscal transfers must hold the OCA together as a country.

The limitations to this approach should be noted. First, the OCA theory would also suggest usage of variables such as fiscal transfers. It is not clear how these variables should be incorporated into the analysis, even if they were readily available by state. Second, the clustering method used is “hard”, in the sense that all observations are allocated to a cluster - perhaps “fuzzy” clustering algorithms (see Bezdek (1974)) may better indicate where cluster outliers could possibly be allocated - giving a better sense of where the borders of the clusters are, and the characteristics that cause these observations to be borderline.

References

Anderberg, M. (1993), *Cluster Analysis for Applications*, Academic Press, New York, USA.

Anderson, E. (1935), "The irises of the Gaspe peninsula", *Bulletin of the American Iris Society* No 59, pp2-5.

Artis, M. and Zhang, W. (1997a), "International Business Cycles and the ERM: Is There a European Business Cycle?", *International Journal of Finance and Economics*, 2, January, pp1-16.

Artis, M. and Zhang, W. (1997b), "On Identifying the Core of EMU: An Exploration of Some Empirical Criteria", *CEPR Discussion Paper* 1689, CEPR, London, UK.

Artis, M. and Zhang, W. (1998a), "Core and Periphery in EMU: A Cluster Analysis", *EUI Working Paper RSC No.* 98/37.

Artis, M. and Zhang, W. (1998b), "Membership of EMU: A Fuzzy Clustering Analysis of Alternative Criteria", *EUI Working Paper RSC No.* 98/52.

Artis, M., Krolzig, H-M., and Toro, J. (1999), "The European Business Cycle", *CEPR Discussion Paper* 2242, CEPR, London, UK.

Banfield, J. and Raftery, A. (1993), "Model-based Gaussian and non-Gaussian clustering", *Biometrics*, Vol. 49, pp803-821.

Baxter, M. and Stockman, A. (1989), "Business Cycles and the Exchange-Rate Regime", *Journal of Monetary Economics*, Vol. 27, pp377-400.

Bayoumi, T. (1994), "A Formal Model of Optimum Currency Areas", *IMF Staff Papers*, Vol.41, No.4, pp537-554.

Bayoumi, T. and Eichengreen, B. (1994a), "One Money or Many?: Analyzing the Prospects for Monetary Unification in Various Parts of the World", *Princeton Studies in International Finance* 76, Princeton University, USA.

Bayoumi, T. and Eichengreen, B. (1994), "Monetary and Exchange Rate Arrangements for NAFTA", *Journal of Development Economics*, Vol 43, pp125-165.

Bayoumi, T. and Masson, P. (1994), "Fiscal Flows in the United States and Canada: Lessons for Monetary Union in Europe", CEPR Discussion Paper 1057, CEPR, London, UK.

Bensmail, H., Celeux, G., Raftery, A. and Robert, C. (1997), "Inference in Model-Based Cluster Analysis", *Statistics and Computing*, Vol 7, pp1-10.

Bezdek, J. (1974), “Cluster Validity with Fuzzy Sets”, *Journal of Cybernetics*, Vol. 3, No.3, pp58-73.

Blanchard, O. and Quah, D. (1989), "The dynamic effects of aggregate demand and supply disturbances", *American Economic Review* 79, pp655-673.

Bock, H. (1996), “Probability models and hypothesis testing in partitioning cluster analysis”, in P. Arabie, L. Hubert, and G. DeSorte (eds), *Clustering and Classification,* World Science Publishers, River Edge, NJ, pp377-453.

Buiter, W. (1998), “The Economic Case for Monetary Union in the European Union”, *Review of International Economics*,

Buiter, W., Corsetti, G. and Roubini, N. (1993), "Excessive Deficits: Sense and Nonsense in the Treaty of Maastricht", *Economic Policy* 16, pp58-60.

Celeux, G, and Govaert, G. (1995), “Gaussian Parsimonious Clustering Models”, *Pattern Recognition*, Vol. 28, No. 5, pp781-793.

Chernoff, H. (1973), “The use of faces to represent Points in k-Dimensional Space Graphically”, *Journal of the American Statistical Association*, Vol 68, pp361-368.

Crone, T. (1999), “Using State Indexes to Define Economic Regions in the US”, Working Paper 99-19, Federal Reserve Bank of Philadelphia, USA.

Crouz, C., Forni, M. and Reichlin, L. (1999), “A Measure of Co-movement for Economic Variables: Theory and Empirics”, *CEPR Discussion Paper* 2339, London, UK.

Crowley, P. (1996), "EMU, Maastricht and the 1996 IGC", *Contemporary Economic Policy*, Vol. XIV, No.2, pp41-55.

Crowley, P. (2000), “Is Canada an Optimal Currency Area? An Exploration of Some Empirical Criteria using Provincial Data”, Paper presented at Western Economic Association meetings in Vancouver, BC, and currently under revision.

Crowley, P. (2001), “Does NAFTA need a single currency?”, Unpublished manuscript, presented at the Western Economic Association meetings in San Francisco, CA.

De Grauwe, P. and Vanhaverbeke, P. (1993), “Is Europe an Optimal Currency Area: Evidence from Regional Data”, in Masson, P. and Taylor, M. (eds.), *Policy Issues in the Operation of Currency Unions,* Cambridge University Press, Cambridge, UK, pp 111-129.

De Serres, A. and Lalonde, R. (1994), “Symétrie des chocs touchant les régions canadiennes et choix d’un régime de change”, *Document de travail* 94-9, Banque du Canada.

Eichengreen, B. (1993), "European Monetary Unification", *Journal of Economic Literature* 31, pp1321-57.

Fisher, R. (1936), “The Use of Multiple Measurements in Taxonomic Problems”, *Annals of Eugenics (London)*, no7, pp179-188.

Fleury, B., Schmid, M. and Narayanan, A. (1994), “Error rates in quadratic discrimination with constraints on the covariance matrices”, *Journal of Classification*, Vol 11, pp101-120.

Fraley, C., and Raftery, A. (1998a), “How Many Clusters? Which Clustering Method? Answers via Model-Based Cluster Analysis”, Technical Report No 329, Department of Statistics, University of Washington, WA, USA.

Fraley, C., and Raftery, A. (1998b), “MCLUST: Software for Model-Based Cluster and Discriminant Analysis”, Technical Report No 342, Department of Statistics, University of Washington, WA, USA.

Fraley, C. and Raftery, A. (2000), “Model-based Clustering, Discriminant Analysis, and Density Estimation”, Technical Report No 380, Department of Statistics, University of Washington, WA, USA.

Frankel, J. and Rose, A. (1997), “Is EMU more justifiable ex post than ex ante?”, *European Economic Review* 41, pp753-760.

Fratianni, M., von Hagen, J. and Waller, C. (1992), *The Maastricht Way to EMU*, Essays in International Finance 187, Princeton University, New Jersey, USA.

Friedman, H. and Rubin, J. (1967), “On some invariant criteria for grouping data”, *Journal of the American Statistical Association*, Vol 62, pp1159-1178.

Galbraith, J. and Jiaqing, L. (1999), “Cluster and Discriminant Analysis on Time-Series as a Research Tool”, UTIP Working paper 6, LBJ School of Public Affairs, University of Texas at Austin, USA.

Gerlach, H. (1988), “World Business Cycles under Fixed and Flexible Exchange Rates”, *Journal of Money, Credit, and Banking*, Vol 20, No 4, pp621-632.

Hartigan, J. (1975), *Clustering algorithms*, Wiley, NY, USA.

Honohan, P. (2000), “Miniblocs and Fringe Currencies of the EMU”, *Journal of Economic Integration*, Vol 15, No 1, pp47-75.

Imbs, J. (1999), “Co-fluctuations”, *CEPR Discussion Paper* 2267, London, UK.

Jacquemin, A. and Sapir, A. (1995), “Is a European Hard Core Credible? A Statistical Analysis”, *CEPR Discussion Paper* 1242.

Kass, R. and Raftery, A (1995), “Bayes factors”, *Journal of the American Statistical Association*, Vol 90, pp773-795.

Kaufman, L. and Rousseeuw, P. (1990), *Finding Groups in Data: An Introduction to Cluster Analysis*, Wiley, New York, USA.

Kenen, P. (1969), “The Theory of Optimal Currency Areas: An Eclectic View”, in (eds), Mundell, R. and A. Swoboda, *Monetary Problems of the International Economy*, University of Chicago Press, Chicago, USA.

LaFrance, R. and St-Amant, P. (1999), “Optimum Currency Areas: A Reviews of the Recent Literature”, Paper presented at the Canadian Economics Association meetings, Toronto, Ontario, May 28-30.

Lalonde, R. and St-Amant, P. (1995), “Zones monétaires optimales: le cas du Mexique et des Etats Unis”, *Document du travail* 93-12, Banque du Canada, Ottawa, Canada

Leonard, H. and Walder, J. (1999), "The Federal Budget and the States, FY1999",

Maharaj, E. and Inder, B. (1999), “Forecasting Time Series from Clusters”, Working Paper 6/99, Department of Econometrics and Business Statistics, Monash University, Australia.

McKinnon, R. (1963), “Optimum Currency Areas”, *American Economic Review*, No. 53, pp717-55.

McLachlan, G. and Krishnan, T. (1997), *The EM Algorithm and Extensions*, Wiley, New York, NY, USA.

Melitz, J. and Zumer, F. (1998), “Regional Redistribution and Stabilization by the Centre in Canada, France, the United Kingdom and the United States: New Estemates Based on Panel Data Econometrics”, CEPR Discussion Paper 1829, London, UK

Mukerjee, S., Feigelson, E., Babu, G., Murtagh, F., Fraley, C. and Raftery, A. (1998), “Three types of gamma ray bursts”, *The Astrophysical Journal*, No 508, pp314-327.

Mundell, R. (1961), “A Theory of Optimum Currency Areas”, *American Economic Review*, Vol. 51, pp657-675.

Murtagh, F. and Raftery, A. (1984), “Fitting straight lines to point patterns”, *Pattern Recognition*, Vol 17, pp479-483.

Neumayer, P. (1998), “Currencies and the Allocation of Risk: The Welfare Effects of a Monetary Union”, *American Economic Review* Vol. 88, pp691-98.

Nilsson, R. (1987), “OECD Leading Indicators”, *OECD Economic Studies*, No 9, OECD, Paris, France.

van Ooyen, A. (2001), “Theoretical Aspects of Pattern Analysis”, in (eds) Dijkshoorn, L, Tower, K. and M. Struelens, *New Approaches for the Generation and Analysis of Microbial Fingerprints*, Elsevier, Amsterdam, Netherlands.

Pedersen, T. (2000), “The Hodrick-Prescott filter, the Slutzky effect, and the distortionary effects of filters”, *Journal of Economic Dynamics and Control*, forthcoming

Sarle, W. (1996), “The Number of Clusters”, SAS Institute, Cary, USA

Scott, A. and Symons, M. (1971) “Clustering Methods based on likelihood ratio criteria”, *Biometrics*, Vol 27, pp387-397.

Struyf, A., Hubert, M. and Rousseeuw, P. (1997), “Integrating Robust Clustering Techniques in S-PLUS”, *Computational Statistics and Data Analysis*, Vol 26, pp17-27.

Tavlas, G. (1993), “The ‘New’ Theory of Optimum Currency Areas”, *The World Economy*, Vol. 16, pp663-85.

Venables, W., and Ripley, B. (1999), *Modern Applied Statistics with S-PLUS*, 3rd edition, Springer-Verlag, New York, NY, USA.

Ward, J. (1963), “Hierarchical groupings to optimize an objective function”, *Journal of the American Statistical Association*, Vol 58, pp234-244.

Wishart, D. (1999), “ClustanGraphics Primer”, Clustan Ltd., Edinburgh, UK

Annex A

Labelling conventions used

AL = Alabama
AK = Alaska
AZ = Arizona
AR = Arkansas
CA = California
CO = Colorado
CN = Connecticutt
DE = Delaware
FL = Florida
GE = Georgia
HI = Hawaii
ID = Idaho
IL = Illinois
IN = Indiana
IO = Iowa
KA = Kansas
KN = Kentucky
LO = Louisiana
ME = Maine
MD = Maryland
MA = Massachusetts
MI = Michigan
MN = Minnesota
MS = Mississippi
MO = Missouri
MT = Montana
NE = Nebraska
NV = Nevada
NH = New Hampshire
NJ = New Jersey
NM = New Mexico
NY = New York
NC = North Carolina
ND = North Dakota
OH = Ohio
OK = Oklahoma
OR = Oregon
PA = Pennsylvania
RI = Rhode Island
SC = South Carolina
SD = South Dakota
TN = Tennessee

TX = Texas
UT = Utah
VT = Vermont
VI = Virginia
WA = Washington
WV = West Virginia
WI = Wisconsin
WY = Wyoming

Annex C

Data Sources

Data Sources	United States
Gross Domestic Product	Bureau of Economic Analysis, U.S. Department of Commerce. http://www.bea.doc.gov/bea/regional/gsp/action.cfm
Unemployment	Bureau of Labor Statistics
Consumer Price Index	See Herman B. Leonard and Jay H. Walder, "The Federal Budget and the States, FY1999," Appendix B for a discussion of the methodology used in constructing these indices.

Annex D

Correlation details

Variable	US	Mean	Standard deviation
Cyclical component of real GDP	1982-1997	0.509	0.396
Unemployment	1978-2000	0.793	0.178
Inflation	1980-1999	0.939	0.132