Chapter 5 Methodology

The previous chapter has described the elements which are specifically needed to meet the first and the second research objective. Of course, they are also used for the rest of the objectives.

Using the monetary anchors and the sets of criteria explained in the preceding chapter, this chapter reveals the ways by which the research objectives can be met. Amongst others, the approach used to draw relatively symmetrical subsets from the set of East Asian countries reviewed is explained. The approach of implementing OCA theory with cluster analysis most probably first used by Michael Artis and Wenda Zhang (Artis & Zhang, 1997a, 2001, 2002) is adopted here along with some modifications to fit the present context.

The specific research questions and objectives are once again reproduced, as displayed in Table 5.1. The discussions in this chapter do matter to the whole research but they cater especially for objectives three to nine as highlighted in the table. The attainment of the objectives is expected to provide the solutions for the respective research questions.

Figure 5.1 depicts the structure of this chapter. Section 5.1 introduces the pattern recognition techniques mentioned in objective three. Criteria weighting exercise using OCA criteria for objective four is discussed in Section 5.2. The approach used to assess level of preparedness for exchange rate unification and for monetary union for objective five is presented in Section 5.3. Subsidiary analysis on OCA criteria for objectives six and seven is explained in Section 5.4. Besides discussing the preliminary analysis and dataset, Section 5.5 also explains the time periods sampled to attain objective eight. Section 5.6 briefs the data analysis process which also explains the way to meet the last objective, objective nine. Section 5.7 concludes.

Table 5.1 Research questions and objectives

	Specific Research Question	Specific Research Objective	
1	How would the grouping configuration differ under different monetary anchor?	To evaluate and compare the results when different monetary anchors, namely dollar, currency basket, yen, euro, and yuan anchors are alternatively assigned.	Cluster analysis methods; Alternative reference countries
2	How different are the partitions when different sets of criteria are used?	To explore and compare the results by OCA with those by Maastricht criteria.	Cluster analysis methods; OCA and Maastricht criteria
3	How would the results differ across different clustering methods?	To assess and compare the results by hierarchical, fuzzy, and model-based cluster analysis methods. Results are also compared with those of principal component analysis.	Cluster analysis methods
4	How would the arrangements vary if benefits and costs of monetary integration are treated equally?	To inspect and compare the solutions when the sum of 'benefit' OCA criteria and the sum of 'cost' OCA criteria are weighted equally.	Cluster analysis methods; Weighted criteria
5	How prepared are generated country clusters for exchange rate fixation and for monetary union?	To infer the degree of readiness for fixed exchange rate and for monetary union by evaluating the groupings of East Asian countries with dollarized and euroized countries respectively.	Cluster analysis methods; Insertion of benchmark cases
6	How dominant are some criteria in representing the rest of the criteria?	To detect and examine subsets of OCA criteria which are most representative of the rest in generating the results.	Cluster analysis methods; Subsets of criteria
7	How important are certain criteria in producing the best partitions?	To detect and assess subsets of OCA criteria which produce the most data-fitting partitions as indicated by particular statistical measures.	Cluster analysis methods; Subsets of criteria; Validation indexes
8	How would the results vary over different economic periods?	To compare the results across pre-crisis, crisis, and post-crisis periods.	Cluster analysis methods; Economic periods
9	How do the findings compare with the actual HongKong-Macau and Singapore-Brunei fixed exchange rate arrangements?	To evaluate the results against the existing fixed exchange rate arrangements of HongKong-Macau and Singapore-Brunei.	Cluster analysis methods; Cluster features

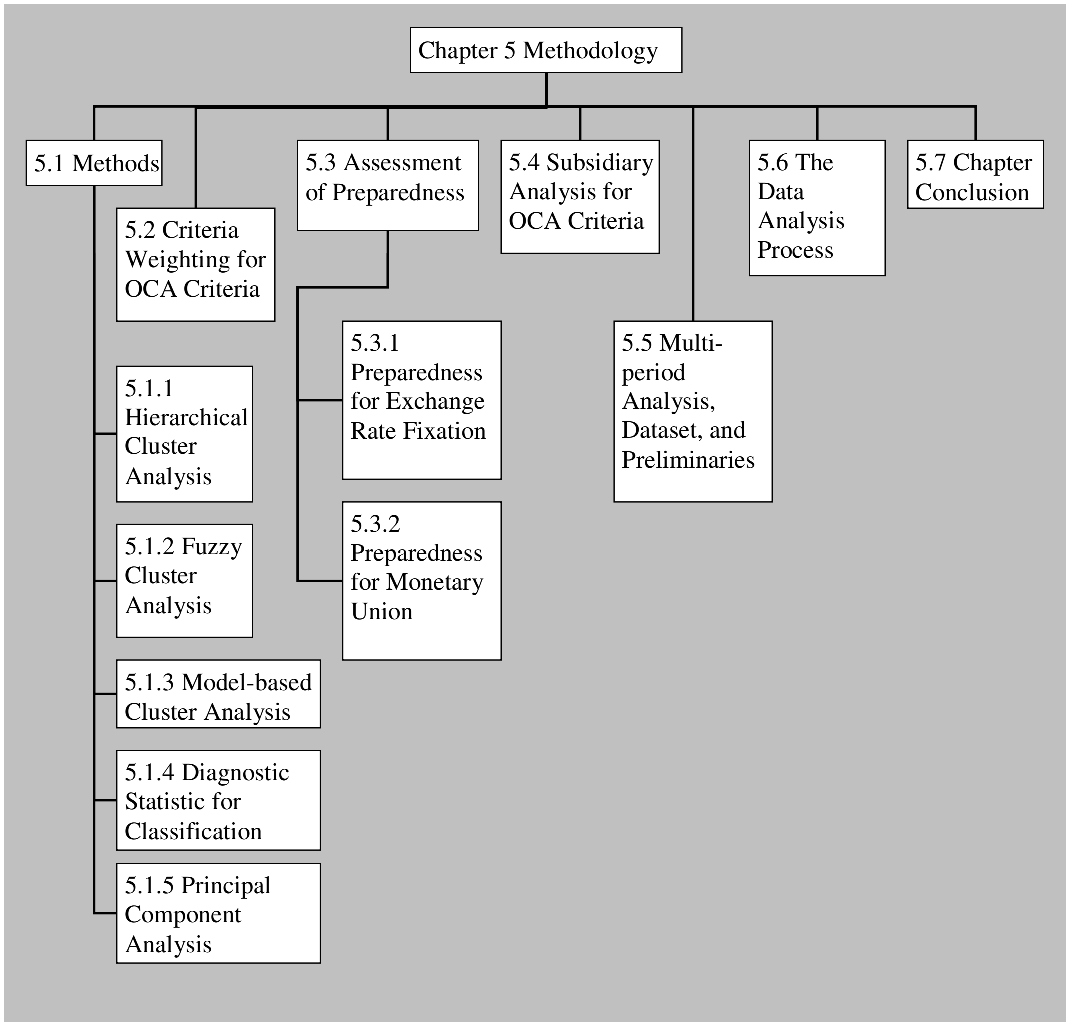


Figure 5.1 Structure of Chapter Five

5.1 Methodology

The dimensions described in the previous chapter may indicate the degree of desirability of fixing the exchange rates for the economies under review. Nevertheless, based on those criteria alone, one cannot extract countries which are sufficiently homogeneous in respect of those facets, namely countries which are equally desirable or relatively fit to constitute a monetary or currency bloc. For this reason, cluster analysis is utilized to suggest comparatively symmetrical subsets of economies given a larger set of countries. Symmetric clusters of countries are important because regardless whether the convergence dimensions are achieved ex-post (see Frankel & Rose, 1998), the countries would be much parallel to begin with, and hence asymmetric experiences that would jeopardize stability of integration will be less likely or less severe.

In addition, cluster analysis also offers the following advantages of which some have already been mentioned in earlier chapters (see also e.g. Anderberg, 1993; Banfield & Raftery, 1993). First, by examining a number of variables simultaneously, cluster analysis enables one to investigate the degrees of symmetry of countries with respect to several relevant dimensions. Hence, each dimension is given equal importance. Second, cluster analysis, a predominantly descriptive technique requires less stringent data requirements (e.g. distribution, stationarity, etc.) than those of conventional econometric approach and thus works well for variables (e.g. the labor criterion) and countries (e.g. the least developed Asian economies) with limited data series. Third, by examining the features of homogenous country groups provided by cluster analysis, one can single out the areas in which a candidate country could improve to achieve convergence with the others. Lastly, hierarchical cluster analysis allows one to explore the evolution of convergence among prospective countries while fuzzy cluster analysis enables one to assess the degree of belongingness of a candidate country to a group of similar countries.

There are some qualifications with cluster analysis though. Cluster analysis does not involve hypothesis testing as opposed to conventional econometrics modeling. But then again since the OCA theory is very difficult to operationalize, any empirical methods that are used obviously explore selected sets of data to try and operationalize the theory. The essence of cluster analysis is therefore to find groupings in a relative sense rather than to assess whether any specific cluster is an OCA in a theoretical sense.

According to Galbraith and Jiaqing (1999), cluster analysis was first used by Fisher (1936) on classifications of irises indigenous to the Gaspe Peninsula in Quebec. The use of cluster analysis on dated information is well-established in disciplines such as geology, paleontology, archeology, and even in biology and developmental psychology. For instance, Chiodi (1989) used time-series height and arm span data to classify children while Hirsch and DuBois (1991) classified children based on the similarities in behavior through time. Infiltration into the economics and finance profession can be seen with the works by Artis and Zhang (1997a), Galbraith and Jiaqing (1999), Maharaj and Inder (1999), Honohan (2000), and recently Artis and Okubo (2009).

In the field of currency areas, empirical studies that have employed cluster analysis grounded in OCA theory include those of Artis & Zhang (2001, 2002), Crowley (2002, 2004, 2008), Bénassy-Quéré & Coupet (2005), Nguyen (2007), Ibrahim (2008), and Tsangarides and Qureshi (2008). These studies have already been discussed in the literature review.

The following sections provide a concise description on the cluster analysis methods employed in this study, namely hierarchical cluster analysis, fuzzy cluster analysis, and model-based cluster analysis. Due to differences in the algorithms, the solutions produced by these techniques are most likely different. For this reason, the solutions are compared and contrasted in the results, as spelt out in the third research objective.

The explanation on hierarchical and fuzzy clustering methods is largely excerpted from Artis and Zhang (2001, 2002) whilst that on model-based cluster analysis is mainly from Crowley (2004).

5.1.1 Hierarchical Cluster Analysis

According to Lorr (1983), hierarchical cluster analysis methods (HCM) are often preferred for classification as it reflects a developmental or evolutionary pattern or sequence. Due to this useful feature, this approach is implemented first in the analysis, followed by fuzzy and model-based cluster analysis methods. The analysis is run using Matlab and the tools provided by Martinez and Martinez (2005).

In the terminology of cluster analysis, there are n objects (cases, observations, countries, etc.) and p variables (features, criteria, dimensions, etc.) in a dataset with each object being denoted by a vector x_i ($x_i = (x_{i1}, x_{i2}, \ldots, x_{ip})$ for $i = 1, 2, \ldots, n$). Each variable is standardized with mean and standard deviation being equal to zero and unity respectively so that they are treated as having equal importance in determining the structure. The same also applies to fuzzy cluster analysis, model-based cluster analysis, and principal component analysis used in this study.

The dissimilarity coefficient or distance d_{ij}, between two objects x_i and x_j is defined by the Euclidean distance:

$$d_{ij} = \sqrt{\sum_{l=1}^{p} (x_{il} - x_{jl})^2} \tag{5.1}$$

The definition of the distance between two clusters is important in determining the shape of homogeneous groups. For hierarchical cluster analysis, there exist few agglomerative algorithms which differ only in the definition of distance between clusters. For details, see Kaufman and Rousseeuw (1990) and Anderberg (1993). Three of the most often used algorithms, namely the average linkage, centroid linkage, and Ward's linkage methods are used in this study. These methods tend to produce spherical clusters.

For average linkage, the distance DistA between two clusters r and s is defined as:

$$\text{DistA} = \frac{1}{n_r n_s} \sum_{i \in r} \sum_{j \in s} d_{ij} \tag{5.2}$$

where n_r and n_s denote the number of objects in clusters r and s respectively. This method tends to combine clusters that have small and approximately equal variances.

Meantime, the centroid linkage defines the distance DistC between two clusters r and s as the Euclidean distance (Eq. 5.1) between their cluster centroids.

A centroid (vector) $\bar{x}(r)$ together with its coordinates $\bar{x}_l(r)$ (for l=1,2,…p), may be expressed as:

$$\bar{x}(r) = (\bar{x}_1(r), \bar{x}_2(r), \ldots, \bar{x}_p(r)) \tag{5.3}$$

where

$$\bar{x}_l(r) = \frac{1}{n_r} \sum_{i \in r} x_{il} \quad \text{for} \quad l = 1, 2, \ldots, p \tag{5.4}$$

A problem with centroid linkage is the possibility of reversals. This can happen when the distance between one pair of cluster centroids is less than the distance between the centroid of another pair that was merged earlier. In other words, the distances between clusters are not monotonically increasing. This could make results confusing and difficult to interpret. When this happens in the results, solutions from centroid linkage are subordinated.

For Ward's linkage, the fusion of two clusters is determined by the size of the incremental sum of squares. It looks at the increase in the total within-group sum of squares when clusters r and s are joined. The distance DistW between clusters r and s is given by:

$$\text{DistW} = n_r n_s DistC^2 / (n_r + n_s) \tag{5.5}$$

Ward's method tends to combine clusters that have a small number of observations. It also has a tendency to locate clusters that are spherical and of the same size. Due to the sum of squares criterion, it is sensitive to the presence of outliers in the dataset.

Each of the above methods starts from a classification with n clusters in it where each cluster contains only one object. The algorithms proceed by successively merging two clusters into one at each stage until a single cluster is obtained. The merging criterion at each stage is to choose two clusters which have the least distance between them. A new classification is identified after two clusters have been merged and the distances between clusters are updated.

Since the agglomerative algorithms differ in their definition of distance, cophenetic correlation coefficient is used to determine the linkage method which best represents the data structure. It is a measure which determines how well the generated clusters represent dissimilarities between objects where values close to 1 representing better clustering. The coefficient measures the correlation between the distances generated by the linkage method and the Euclidean distances between the objects.

Letting *d* be the average of d_{ij} and letting *t* be the average of the t_{ij}, the distance generated by a linkage method at which two objects x_i and x_j are first joined together; the cophenetic correlation coefficient *r* (*r* here stands for correlation coefficient) is given by:

$$r = \frac{\sum_{i<j}(d_{ij}-d)(t_{ij}-t)}{\sqrt{(\sum_{i<j}(d_{ij}-d)^2)(\sum_{i<j}(t_{ij}-t)^2)}} \tag{5.6}$$

The outcome of hierarchical clustering is presented in the form of a tree known as dendrogram. The heights of the links of the dendrogram represent the distance at which each fusion is made such that greater dissimilarity between objects is reflected by larger distances and taller links.

While the dendrogram provides some indication on the number of clusters, the 'optimal' number however could be rather subjective depending on the observer. In light of this, the pseudo-F index or Calinski-Harabasz index (CHI)[44] developed by Calinski and Harabasz (1974) is used. Indeed, it has been detected by Milligan and Cooper (1985) to be the best measure among thirty cluster-stopping rules. This index has been used to determine optimal number of clusters by recent hierarchical clustering studies by Tsangarides and Qureshi (2008) and Artis and Okubo (2009) and is defined as:

[44] Note: "CHI" in this case is not "Chi-square".

$$\mathrm{CHI} = \frac{S_b\,/(k-1)}{S_w\,/(n-k)} \tag{5.7}$$

where S_b is the between-cluster sum of squares, S_w is the within-cluster sum of squares, k is the number of clusters, and n is the number of objects. Higher index values signify more distinctive partitioning and better clustering.

To complement the CHI, another stopping rule is also used. Different from CHI in which higher values suggest better partitioning, for C-index lower values are associated with better classifications. The C-index (Hubert & Schultz, 1976) is defined as:

$$\mathrm{C} = \frac{S - S_{\min}}{S_{\max} - S_{\min}} \tag{5.8}$$

in which S is the sum of distances over all pairs of objects from the same cluster whereby q is the number of those pairs and S_{min} is the sum of the q smallest distances when all pairs of objects are considered. Likewise, S_{max} is the sum of the q largest distances out of all pairs. The C-index should be minimized and it is limited to the interval [0, 1]. In the findings, the indications by CHI and C-index are fairly consistent.

5.1.2 Fuzzy Cluster Analysis

While hard clustering approach (e.g. hierarchical method) attempts to assign each object to one and only one cluster or group, fuzzy clustering method (FCM) allows some ambiguity in the data by assigning each object to a cluster with a membership coefficient indicating the degree of belongingness of the object to that cluster (see Dunn 1973; Bezdek, 1973; Hoppner, Klawonn, Kruse, & Runkler, 1999). In a sense, it has more power in approximating the situation involving incomplete and uncertain information, which is often the case in the real world. An object is most likely to belong to the cluster with which it has the highest membership coefficient.

The algorithm of fuzzy analysis used in the analysis is the widely used fuzzy C-means technique proposed by Dunn (1973) and Bezdek (1973). The algorithm is briefly explained here. Similar to HCM, the FCM algorithm computes with the standard Euclidean distance norm which induces spherical clusters. Hence, it can only detect clusters of the same shape and orientation. The algorithm is based on the minimization of the following objective function:

$$J(u,v) = \sum_{i=1}^{n} \sum_{k=1}^{c} (u_{ik})^2 d_{ik}^2 \tag{5.9}$$

subject to the following constraints:

$$0 \le u_{ik} \le 1, \quad \sum_{k=1}^{c} u_{ik} = 1, \quad 0 < \sum_{i=1}^{n} u_{ik} < n \quad \text{for } i = 1,2,...,n; \ k = 1,2,...,c \tag{5.10}$$

in which u_{ik} stands for the membership coefficient of object x_i belonging to cluster k and c is the number of clusters. d_{ik} is the Euclidean distance between vector x_i and the center of the cluster k, a p-dimensional vector v_k where:

$$v_k = \sum_{i=1}^{n} u_{ik}^2 x_i \Big/ \sum_{i=1}^{n} u_{ik}^2 \tag{5.11}$$

The above algorithm is based on the assumption that the number of clusters is known in advance. In reality, however, researchers have to choose the number of clusters so as to ensure that the clusters are as ‘crisp’ as possible. Hence, to determine the optimal number of clusters, Xie and Beni’s (1991) index (XBI) is used here:

$$\text{XBI} = \frac{1}{n} \frac{\sum_{i=1}^{n} \sum_{k=1}^{c} (u_{ik})^2 d_{ik}^2}{\min_{i,k} d_{ik}^2} \tag{5.12}$$

Recent application of this index with FCM includes the works by Nguyen (2007) and Tsangarides and Qureshi (2008). In fact, it is one of the most frequently used indexes for cluster analysis involving fuzzy memberships (see Fontán & Jiménez, 2004). Low indexes indicate less (greater) variations within (between) clusters. Hence, smaller index values represent more compact and separated clusters.

With respect to the present study and data, XBI has virtually provided unique solution, that is, only one minimum value virtually all the time when compared to other compatible measures mentioned in Balasko, Abonyi, and Feil (2004).[45] This finding further justifies the use of this measure.

There has also been another diagnostic statistic in fuzzy analysis, Dunn's Partition Coefficient (DPC) which measures the degree of fuzziness in the partitions (see Dunn, 1973). DPC is defined as the sum of squares of all the membership coefficients divided by the number of objects and may be further normalized as in the following formula:

$$\mathrm{DPC} = \frac{c\sum_{i=1}^{n}\sum_{k=1}^{c}\frac{u_{ik}^{2}}{n} - 1}{c-1} \tag{5.13}$$

The normalized DPC, varying from 1 to 0 is a useful indicator of the data structure; a value close to 1 indicates no fuzziness in the data whilst a value close to 0 indicates complete fuzziness.

For this study, cluster solutions are selected based on XBI provided that the corresponding DPC is reasonably high to enable meaningful interpretations of the partitions. FCM is run using Matlab with the aid from Fuzzy Clustering and Data Analysis toolbox by Balasko, Abonyi, and Feil (2004). More information on the fuzzy algorithm and the scalar validity measures used here can be found in the toolbox manual. For greater details, such as on determination of the membership coefficients, see an extract from the manual is provided in Appendix B.

5.1.3 Model-based Cluster Analysis

Besides HCM and FCM, there is a relatively new clustering method, model-based clustering (MBC) which was first used by Patrick Crowley in the OCA context (see Crowley, 2002, 2004, 2008). The MBC method is based on probability models, such as

[45] Other measures that can be used with the fuzzy algorithm include Partition Coefficient, Classification Entropy, Partition Index, and Separation Index.

the finite mixture model for probability densities (see Fraley & Raftery, 1998).

According to Crowley (2004), the alternative of MBC arises in the midst of weaknesses in conventional clustering methods wherein the number of clusters must either be specified first or chosen later based on some validation indexes (see e.g. Bock, 1996; Martinez & Martinez, 2005). Nonetheless, up till recently none of these methods have been satisfactory from a computational or methodological point of view. Another problem is that those clustering methods impose a certain structure on the clusters (usually spherical) and the statistical properties of the clusters are generally unknown. Quite the reverse, MBC is capable of identifying more flexible cluster structures such as ellipsoidal clusters. In light of this, it is not surprising that configurations generated by MBC could be very different from those produced by other methods.

Though MBC is advantaged in terms of computational rigor, from the perspective of application, HCM and FCM could be more useful in the sense that they provide users such as policymakers with information which can assist them in their decision-making. To reiterate, HCM could help in the sequencing process of potential countries while FCM could indicate how close a country is in relation to each of the identified homogenous groups of countries.

The following discussion briefly explains the MBC methodology. In probability based clustering, each observation x_i is assumed to be generated by a mixture of underlying probability distributions where each component in the mixture represents different cluster. Given a set of observations, the density of an observation x_i from the k-th component in a total number of c components is $f_k(x_i|\theta_k)$ where θ_k are the parameters. In most cases, $f_k(x_i|\theta_k)$ is assumed to be multivariate normal (Gaussian), so in this instance the parameters θ_k consist of a p-dimensional mean vector μ_k and a $p \times p$ covariance matrix Σ_k. The clusters will then be ellipsoidal with center at μ_k and the covariance matrix Σ_k will determine the other characteristics. The mixture likelihood approach then maximizes the criterion:

$$\ell_M\ (\theta_1,...,\theta_c : \pi_1,...,\pi_c \big| x_i) = \prod_{i=1}^{n} \sum_{k=1}^{c} \pi_k f_k (x_i \big| \theta_k) \tag{5.14}$$

where π_k is the probability that an observation belongs to the k-th component.

Banfield and Raftery (1993) and Celeux and Govaert (1995) developed a model-based framework for clustering by expressing the covariance matrix in terms of its eigenvalue decomposition which is of the form;

$$\Sigma_k = \lambda_k D_k A_k D_k^T \tag{5.15}$$

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 Σ_k, and λ_k is a scalar. This leads to a geometric interpretation of clusters; D_k determines the orientation, A_k determines the shape of the density contours, and λ_k specifies the volume. These characteristics can be allowed to vary between clusters, or constrained to be the same for all clusters. This approach actually subsumes many previous approaches of model-based clustering. The parameterizations of covariance matrix for different models are displayed in the following table, Table 5.2.

Table 5.2 Parameterizations of covariance matrix by model[1]

M[2]	Covariance	Distribution				Description[3]			
		Family	Volume	Shape	Orientation				
1	$\mathbf{\Sigma_k} = \lambda \mathbf{I}$	Spherical	Equal	Equal	Not Applicable	Diagonal covariance matrices	Diagonal elements are equal	Covariance matrices are equal	**I** is a *p* × *p* identity matrix
2	$\mathbf{\Sigma_k} = \lambda_k \mathbf{I}$		Variable					Covariance matrices may vary	
3	$\mathbf{\Sigma_k} = \lambda \mathbf{B}$	Diagonal	Equal		Axes		Diagonal elements may be unequal	Covariance matrices are equal	**B** is a diagonal matrix
4	$\mathbf{\Sigma_k} = \lambda \mathbf{B}_k$			Variable				Covariance matrices may vary among components	
5	$\mathbf{\Sigma_k} = \lambda_k \mathbf{B}_k$		Variable						

6	$\boldsymbol{\Sigma}_{\mathbf{k}} = \lambda \mathbf{D}\mathbf{A}\mathbf{D}^{\mathbf{T}}$	General	Equal	Equal	Equal	Covariance matrices can have nonzero off-diagonal elements	Covariance matrices are equal
7	$\boldsymbol{\Sigma}_{\mathbf{k}} = \lambda \mathbf{D}_k \mathbf{A}\mathbf{D}_k^T$				Variable		Covariance matrices may vary among components
8	$\boldsymbol{\Sigma}_{\mathbf{k}} = \lambda \mathbf{D}_k \mathbf{A}_k \mathbf{D}_k^T$			Variable			
9	$\boldsymbol{\Sigma}_{\mathbf{k}} = \lambda_k \mathbf{D}_k \mathbf{A}_k \mathbf{D}_k^T$		Variable				

Notes:
1 This is a subset of the available models. 2 *M* for model. 3 *p* is number of dimensions (variables).
Source: Martinez and Martinez (2005).

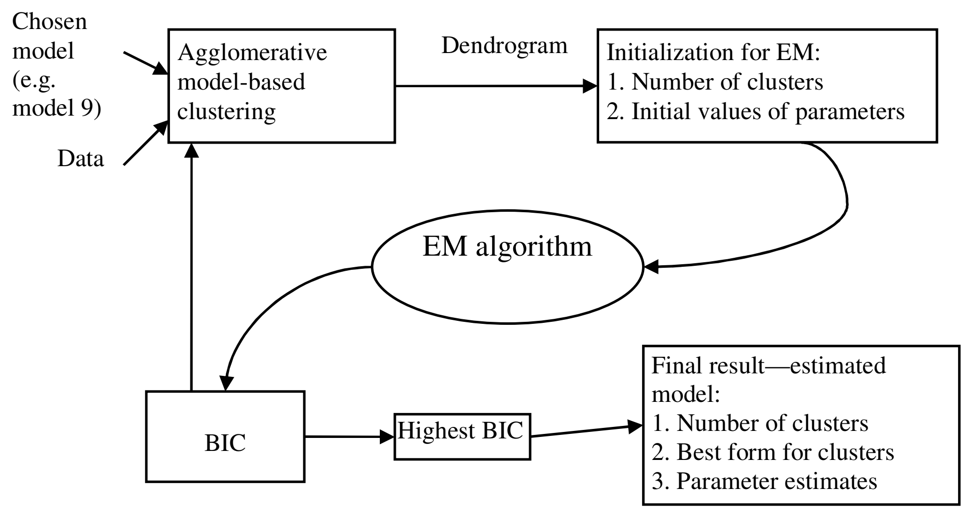
Given the different model parameterizations above, agglomerative hierarchical clustering can be used by merging clusters so as to maximize the resulting likelihood as specified in equation (5.14) above. The algorithm used for maximizing the likelihood function is the Expectation-Maximization) EM algorithm (see Redner &Walker, 1984). EM iterates between an E-Step, which calculates the posterior probability that the i-th observation belongs to the k-th component given the current values of the parameters, and an M-Step which updates the parameter estimates using the estimated posterior probability. In the limit, the parameters usually converge to the maximum likelihood values for the Gaussian mixture model.

The mixture model approach allows the use of approximate Bayes factors to compare the appropriateness of the models. The Bayes factor is the posterior odds for one model against the other assuming neither is favored a priori. A convention is to choose the model and the number of clusters according to the Bayesian Information Criterion (BIC):

$$2\log p(x|M) + \text{const} \approx 2\ell_M(x|\theta) - m_M \log(n) \equiv \text{BIC} \tag{5.16}$$

where $p(x|M)$ is the likelihood of the data for the model M, $\ell_M(x|\theta)$ 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 and hence the cluster solution. A standard convention for calibrating BIC differences is that differences of 10 or more correspond to ‘strong’ evidence.

Figure 5.2 illustrates the MBC process, implemented using Matlab and the Model-based Cluster Analysis toolbox by Martinez and Martinez (2005). Explanation on the steps from Martinez and Martinez is placed in Appendix C. For this study, the number of clusters and hence the partitioning of countries from the results are used.



Source: Martinez and Martinez (2005).

Figure 5.2 The model-based clustering process

5.1.4 Diagnostic Statistic for Classification

Among many cluster validation indexes available, silhouette width (SW) is one conventional statistic used to indicate which objects lie well within their cluster and which ones are merely somewhere in between clusters (see Rousseeuw, 1987). Unlike most other indexes which provide aggregate measures, SW is able to suggest the appropriateness of clustering at individual object level. Given individual silhouettes, the silhouette of a cluster is simply the average of the individual silhouettes of the objects lying in the cluster.

Because of this useful feature, SW is used here to compare the ‘tightness’ of groupings amongst clusters by each of the cluster analysis methods. It is also used to suggest amongst the clustering methods and between the OCA and Maastricht criteria sets, which method and set of criteria are compatible with better partitioning.

The silhouette width of an individual object x_i is defined as:

$$s(i)=\begin{cases}1-a_i/b_i & \text{if } a_i<b_i\,,\\ 0 & \text{if } a_i=b_i\,,\\ b_i/a_i-1 & \text{if } a_i>b_i\,.\end{cases} \tag{5.17}$$

$$\text{or,}\qquad s(i)=\frac{b_i-a_i}{\max[a_i,b_i]}\quad;\quad -1\le s(i)\le 1$$

where a_i is the average distance between object x_i and all other objects in the same cluster and b_i denotes the smallest distance between x_i and other clusters. A value close to 1 indicates that the object is well-classified whereas a value near 0 signals high degree of fuzziness and the object might be better classified to a neighboring cluster. A negative silhouette value indicates that the object is misspecified.

5.1.5 Principal Component Analysis

Principal component analysis (PCA) has been used as a robustness check for cluster analysis solutions by Tsangarides and Qureshi (2008). Concisely, PCA, one of the established dimensionality reduction techniques is a procedure for multivariate analysis which reduces the number of possibly correlated variables into a smaller number of uncorrelated variables known as the principal components while at the same time accounting for as much of the variation in the original dataset as possible.

Each principal component PC, is a linear combination of the original variables which may be expressed as $\mathrm{PC_i}=\alpha_{i1}\mathbf{X}_1+\alpha_{i2}\mathbf{X}_2+\cdots+\alpha_{ip}\mathbf{X}_p$, where i = 1, 2, …, p, and $\mathbf{X}$ is a data matrix with n observations and p variables. The first PC is supposed to account for much of the variability in the data whereas the subsequent PCs explain the remaining variation.

The main purpose of PCA is to gain information and understanding of the data by looking at the observations in the new space. Through a multi-dimensional projection, the distances among objects can be visualized—a feature not offered by any of the clustering methods discussed earlier. By comparing the cluster analysis classification with the PCA projection, one can better understand the ways the cluster analysis methods partition the objects. Unlike the cluster analysis methods however, PCA does not assign objects with cluster memberships.

In this study, the number of components used depends on the *k*th component at which the slope in the scree plot starts to level off markedly, a convention used in PCA (see e.g. Martinez & Martinez, 2005). If that condition is not present, the components should collectively explain at least 60–70 percent of variation in the data. In the results, the first two or three of the components have been found to satisfy either the first or the second condition or both. For more details on PCA and its application with Matlab used for this study, see for instance Jolliffe (2002) and Martinez and Martinez (2005) respectively.

Summarizing the methods:

Hitherto, the main analytical tools have been discussed. In summary, there are 3 clustering methods, that is, HCM, FCM, and MBC. For HCM, there are 3 linkage methods, namely average, centroid, and Ward's linkage methods.

These are the unique features of the methods. HCM provides the merging process as portrayed by the dendrogram; FCM provides degrees of belonging for each object; and MBC offers a more rigorous approach to determination of number of clusters whilst catering for different patterns of clusters as opposed to spherical types by HCM and FCM. HCM and MBC both produce absolute memberships for each object, unlike FCM which provides fuzzy memberships.

Each clustering method will be performed sequentially on the same set of data to enable comparisons for the results across the methods.

5.2 Criteria Weighting for OCA Criteria

The variables used in the analysis discussed above are standardized and weighted equally, hence each criterion is assumed to be of equal significance. Nevertheless, recall that the OCA dimensions discussed in the preceding Chapter 4 can actually be categorized into those measuring the 'benefit' and those representing the 'cost' aspects of exchange rate unification.

To recap, two of the criteria, trade openness and external indebtedness may be more suitably interpreted as measures of benefit. For trade, the more open the economy to the reference country, the higher will be the benefits from exchange rate certainty with that country. For debt liability, the larger the degree of external indebtedness, the greater the value from exchange rate stability with the currency denominating the debt.

The rest of the criteria might more appropriately be regarded as measures of cost wherein the greater the conformity to the preconditions, the weaker will be the likelihood of asymmetric shocks between economies or the smaller will be the adjustment costs when asymmetric shocks do occur. These criteria are real business cycle symmetry, labor market flexibility, export diversification, inflation convergence, real exchange rate stability, and real interest rate cycle symmetry.

Following Artis and Zhang (2001), to avoid the problem of 'costs' being given a bigger weight than the 'benefits', the cluster analysis is also run by weighting the sum of the benefit variables and the sum of the cost variables equally. In Artis and Zhang, inflation convergence is treated as a 'normalization' variable whilst for the reasons explained in the previous chapter, it is viewed as a measure of cost here.

The results using the weighted criteria will be compared to the original results. This exercise will meet research objective four when the respective analyses are done in the findings chapters.

5.3 Assessment of Preparedness

While cluster analysis could produce comparatively symmetrical subsets of countries in the relevant dimensions, an essential condition for integration, it nevertheless could not indicate how prepared are the countries for monetary unification. Indeed, assessment of the level of preparedness for fixed exchange rates and for monetary union is rarely emphasized in clustering-based OCA works.

This section explains how the levels of preparedness amongst the East Asian countries can be inferred. In brief, associations or groupings with countries which maintain rigid exchange rate pegs (i.e. dollarization) and which maintain currency union arrangement (i.e. EMU) are expected to indicate the level of readiness for fixed exchange rates and monetary union respectively amongst the Asian economies. This exercise is done to meet the fifth research objective.

5.3.1 Preparedness for Exchange Rate Fixation

To infer the levels of preparedness for bilateral fixed exchange rate with the reference country among the Asian countries, associations with dollarized countries observed through their groupings in the clustering solutions are used as an indication. If an Asian country shares the same cluster with the dollarized cases, that country could possibly possess characteristics which are essential for exchange rate fixation. The reference country used for the dollarized countries is the US.

Constrained by data availability over the criteria, the selected dollarized countries or benchmark cases are coincidentally Latin American countries, namely Ecuador, El Salvador, Guatemala, and Panama. Panama officially dollarized in 1904, Ecuador in 2000, and El Salvador and Guatemala in 2001 (Castillo, 2006).[46]

[46] Ecuador replaced its sucre with the dollar in September 2000. On Jan. 1, 2001, El Salvador followed suit, and Guatemala elevated the dollar to equal status with its quetzal on May 1. Dollarization refers to the replacement of local currencies with the US dollar in both local and international monetary transactions. Full or official dollarization occurs when a country completely gives up its national currency

In order to account for different levels of preparedness, the dollarized cases, except for Panama which has already adopted the US monetary standard since 1904, are categorized into time periods according to their 'proximity' toward implementation of dollarization. For Ecuador, El Salvador, and Guatemala which have only begun to dollarize in recent decade in 2000 or 2001, the observations for the post-dollarization period 2001–2007 are labeled as ECU3, ELS3, and GTM3 respectively. Groupings with these cases might indicate greater degree of preparedness for fixed exchange rates than with the cases for the earlier periods of 1981–1996[47], the 'pre-crisis' period of which the cases are identified as ECU1, ELS1, and GTM1, and 1997–2000, the 'crisis' period with labels ECU2, ELS2, and GTM2.

While the approach appears to be persuasive, there are several assumptions that need to be clarified. First, it assumes that countries which share similar OCA characteristics with the dollarized countries, that is, sharing the same groupings with them especially the post-dollarization observations in the classification results are relatively more ready to adopt foreign monetary policy and to fix their exchange rates. Undoubtedly, this assertion would be more appropriate when the reference country is the US than when the reference country is other country. In spite of this, there seems to be very few countries, if any, which adopt the monetary standards of other potential anchor countries considered in this research. In fact, except for countries adjacent to the EU which adopt the euro, no countries have actually adopted the Japanese yen (see e.g. Oomes & Meissner, 2008) or the Chinese yuan as official monetary standard. Along these lines, for consistency and comparability purposes, the exercise using dollarized

---

and instead adopts the dollar as its official unit of currency.

[47] In the period before 1998, Latin Americas grew at about 3.5 per cent a year, equivalent to 1.7 per cent per capita, which was slightly above the rate of growth of the rest of the world. During this period, productivity and investment were increasing rapidly. Unfortunately, the Russian crisis came in 1998, and the following year saw a significant retrenchment of capital flows into the region. The most vulnerable countries, Ecuador, Venezuela, and Brazil experienced currency crises, and Argentina began a protracted recession and could not maintain its currency peg. Argentina carried neighbor Uruguay with it. For more, visit http://www.metrocorpcounsel.com/current.php?artType=view&artMonth=March&artYear=2009&-EntryNo=3827 (Retrieved April 9, 2009)

cases is repeated for other monetary anchors as well. If certain Asian economies' OCA features relative to a reference country, say Japan, are similar to the dollarized countries' features against the US, those economies might be more feasible for yen adoption.

Hence, the second assumption is that those OCA features which have 'worked' for the dollarized countries (if they did, with the US dollar as the anchor) would also work for the Asian countries (for other any monetary anchor examined here). Obviously this is only reasonable if the OCA criteria had worked for the dollarized countries in the first place.

The approach of using the Latin American countries as the 'benchmarks' is also logical in some other ways. First, in the exercises when the US is set as the reference country, benchmarking against the dollarized countries is obvious. In fact, the whole Asia Pacific region is closely connected with the US in real and monetary terms (Jameson, 1990; McKinnon, 2005). Second, countries in East Asia and Latin America had been facing similar business cycle experience, predominantly driven by openness to international capital flows (see Rogoff, 2005). For instance, the Asian crisis which began in East Asia actually spilled over to Brazil in the following year and subsequently the whole Latin American continent (see Beckerman & Solimano, 2002).[48] The segmentation of the crisis period 1997–2000 in the analysis does reflect this.

5.3.2 Preparedness for Monetary Union

At most, associations with dollarized countries might only imply the degree of readiness for bilateral fixed exchange rate arrangement. Using the same reasoning, associations with euroized countries might in turn signify the levels of preparedness for multilateral

[48] In 1998 and 1999, majority of Latin American countries suffered a severe recession that drove corporations into insolvency and bankruptcy. The impact of the crisis was particularly violent in the region's smaller economies such as Bolivia, Ecuador, and Uruguay, but countries like Argentina, Colombia, and Chile also suffered severely. Many agree that the major cause for the crisis was the enormous weight taken by short-term financing in emerging economies in Asia and the Pacific as part of the process towards liberalization and globalization.

currency unification.[49] If the features of a group of Asian economies relative to a reference country, say US, are as good as the corresponding features of the EMU founding members against Germany, the de facto center country for EMU, the subset of Asian economies might be relatively prepared than other Asian counterparts for a monetary union anchored on the US dollar.

Whilst the exercises involving the dollarized cases utilize only OCA criteria, it is persuasive to sequentially use the OCA and the Maastricht criteria in the analysis involving euroized benchmarks.

The euroized benchmark cases are computed according to the time periods representing different phases of EMU integration. The periods are the pre-Maastricht period of 1988–1993, the post-Maastricht period of 1994–1998, and the post-euro period of 1999–2007. The data period is sliced in such a way to reflect three milestones in the path toward union: (1) July 1, 1987, the effective date of Single European Act; (2) November 1, 1993, the effective date of Maastricht Treaty; and (3) January 1, 1999, the day the single currency euro standard was launched and implemented since.[50]

Amongst the 11 Western European countries which have officially euroized since

---

[49] In contrast to the associations with the dollarized cases, linkages with EMU cases could yield different inferences due to some distinctions that exist between the European monetary union and dollarization.

Among others; (1) the EMU involves the setting up of supranational authorities governing monetary and perhaps fiscal policies of member countries whereas dollarization involves eschewing national monetary policies and adoption of the US policies; and (2) the EMU involves circulation of a new single common currency whereas dollarization involves partial or full adoption of the US dollar as the/or one of the national money/s. Thus, here the EMU benchmarks are used to signify readiness for EMU-like monetary union whereas the dollarized benchmarks are used to infer readiness for bilateral exchange rate fixation.

[50] Here is a brief account of the landmark events. The Single European Act was the first major revision of the Treaty of Rome that formally established the single European (SEA) market and the European Political Cooperation. The goal of SEA was to remove remaining barriers between countries, to increase harmonization, and to enhance the competitiveness of European countries. It reformed the operating procedures of the institutions and Qualified Majority Voting was extended to new areas.

The Maastricht Treaty (formally, the Treaty on European Union) was signed on February 7, 1992 in Maastricht, the Netherlands after final negotiations on December 9, 1991 between the members of the European Community and entered into force on November 1, 1993 during the Delors Commission. It created the EU and led to the creation of the euro, and is commonly referred to as the pillar structure of the EMU. The pillars are the European Community, the Common Foreign and Security Policy, and the Justice and Home Affairs pillars.

On January 1, 1999, in the first wave of EMU, 11 European countries, Austria, Belgium, Finland, France, Germany, Ireland, Italy, Luxembourg, the Netherlands, Portugal, and Spain adopted the euro, the single common currency.

1999, it is widely accepted that Germany is the de facto center country (see e.g. Artis & Zhang, 2001, 2002; Boreiko, 2003; Crowley 2008). Therefore, in the analysis the observations for euroized cases are computed for 10 founding members only and Germany is used as the reference country for them. For every dimension, average over countries is used and the cases for the pre-Maastricht period, the post-Maastricht period, and the post-euro period are labeled as EMU1, EMU2, and EMU3 respectively.

As the case with dollarized countries, several qualifications need to be laid out. First, it is assumed that the OCA and the Maastricht dimensions have been important for eurozone founders in facilitating the integration process. Second, the eurozone experience is assumed to be replicable in East Asia. Third, associations with post-euro EMU cases are expected to signify higher readiness for common monetary standard whilst those with pre-euro cases are expected to imply lower readiness.

Unquestionably, in many aspects, the foundations in collaboration whether economic or political in Western Europe are remarkably advanced from those in East Asia and hence, conditions that have characterized euro formation might not be generalizable to East Asia. Moreover, an internal anchor was chosen for EMU whereas the appropriate monetary anchor for Asia is still open for question.

Despite the above and the recent crisis in eurozone, the only currency union of sovereign states which has been highly credible to date is none other than EMU and therefore it is compelling to use the EMU as a 'benchmark'. In fact, similar approach comparing the 'readiness' of East Asia to that of EMU is frequently seen in empirical works (e.g. Bayoumi & Eichengreen, 1994; Kawai & Motonishi, 2005; Huang & Guo, 2006).

Due to lack of consistent data over Asian and euroized countries, the OCA external indebtedness criterion used for the Asian cases is not computed for the EMU cases for this preparedness assessment exercise. In fact, this constraint is one reason why indebtedness is not included as one Maastricht criterion in this study because comparisons with the euroized benchmarks would not make sense if Asian cases use different definitions of indebtedness. For eurozone countries, data series for public debt are easily available but not those for private debt. For East Asian countries, only gross external debt stocks which comprise of both public and private external debts are adequately available. Thus, except of this debt variable, exactly the same sets of OCA and Maastricht criteria are used for EMU and Asian cases. The 'missing' debt data for EMU benchmarks are replaced with mean over all countries in the analysis.

5.4 Subsidiary Analysis for OCA Criteria

The following exercises are carried out for all cluster analysis methods using OCA criteria only.

Criteria Dominance

The analyses discussed thus far involve all the OCA dimensions presented in Chapter 4. Nonetheless, it may also be compelling to explore how far concentration on smaller subsets of the OCA criteria is representative of the rest of the criteria. It can be regarded as checking whether certain criteria could be described as dominant (Artis & Zhang, 2001). To implement this exercise, variables are removed one by one using all possible ways of sequencing the variables into the cluster analysis, as long as the number of countries dislodged from their original clusters does not exceed 3. If more than one set of variables satisfies this condition (in the findings, this is extremely rare), the subset of variables which produces the most similar configuration to the original one is preferred. This exercise is performed to attain objective six.

Variable Selection

In addition, it may also be informative to identify subsets of variables which could produce the best partitions by the standards of the respective validation indexes or stopping rules associated with the clustering techniques. The exercise is done by sequentially entering combinations of variables into the analysis and the sets of variables indicated to have generated the best partitions are determined. The analysis will meet objective seven.[51]

5.5 Multi-period Analysis, Dataset, and Preliminaries

Data Period

For the present paper, analysis across non-overlapping successive periods allows the study to assess the stability of country groupings and to explore the pattern of the cluster configuration over periods. This exercise would address the eighth objective. Authors who have used this approach in OCA literature include Font-Vilalta and Costa-Font (2006) and Crowley (2008).

As briefly mentioned in Chapter 1, data analysis is carried out sequentially for three sample periods: the ‘pre-crisis’ or ‘growth’ period of 1981–1996, the ‘crisis’ period of 1997–2000, and the ‘post-crisis’ period of 2001–2007. The rationale for such segmentation is as follows:

- The pre-crisis period 1981–1996 is the period prior to the Asian financial crisis when the region had been experiencing high economic growth—coined by the World Bank as the ‘East Asian Miracle’. Indeed, among the emerging markets, the real GDP from 1965 to 1993 grew at an average annual rate of nearly 9 percent, more than twice as fast as their Latin American counterparts (WB, 1993). Following Font-Vilalta and Costa-Font (2006), the sample period of 1981–1996 also takes into account the structural change after the petroleum crises in 1979.
- The crisis period 1997–2000 is the period of financial and economic distress. The

[51] For more discussion on criteria weighting and selection of variables in cluster analysis, see for instance, Gnanadesikan, Kettenring, and Tsao (1995).

same period has also been examined separately by Font-Vilalta and Costa-Font (2006) and Nguyen (2007) in their OCA analysis on East Asia. It is crucial to examine this period separately to check whether solutions from this period are significantly different from those of other periods. If results are indeed different, the Asian crisis could have great impact on the grouping arrangements. In another aspect, recall that dollarized countries in the Americas are used as the benchmarks for assessing the preparedness of Asian countries for fixed exchange rates. Hence, the period also represents the 1998–1999 Latin American crisis which was partly triggered by the Asian crisis (see Beckerman & Solimano, 2002).[52] Indeed, as shown by the preliminary analysis in the findings chapters later, data patterns for certain variables of the crisis period are significantly different from those of other periods.

- The post-crisis period 2001–2007 is assessed as a distinct period as many believe that the crisis has driven the region toward greater integration and multilateral cooperation (see e.g. Plummer, 2007; Rana, 2007). Some of these initiatives have already been discussed in Chapter 2.

Dataset

Regarding the dataset, monthly series are used except when consistent data are unavailable. Major sources include IMF, ADB, and WB databases in addition to national and central bank statistical websites and databases. Precise data definitions and sources are placed in Appendix A.

While data series used are generally from 1981 to 2007, the interest rate series are nevertheless started from 1992 to ensure the greatest degree of consistency over countries. Meanwhile, the data for the OCA dimension labor market flexibility are only

[52] Though some may argue that the crisis had actually ended earlier, the recovery date however differs somewhat from country to country. Hence, it is better to cover a longer period here.

available for the post-crisis period (equivalently post-dollarization period for dollarized cases or post-euro period for euroized cases).

Hence, for the OCA-based pre-crisis and crisis period analyses, exercises are carried out with 7 variables. For post-crisis period, two sets of results are shown; one using 7 variables excluding the labor criterion and the other using 8 variables including the labor criterion. This will enable direct comparisons over the periods and at the same time for the post-crisis period, results with and without the labor criterion can be compared.

Whenever data are not available or missing for an observation, variable mean over all countries is used. These are the missing data: pre-crisis period interest rate series for Brunei; the labor data for Myanmar and Macau; and the post-dollarization interest rate series for El Salvador. The variable mean is also used for the external debt criterion for the advanced economies of Japan, Australia, New Zealand, Canada, and the EMU cases.

To compute the variables for the currency basket anchor, the G3 weights suggested by Williamson (2005) (US 0.47; Japan 0.23; Germany/EMU 0.30) are used for reference-dependent criteria, that is, variables measured relative to a reference. For OCA, 5 of the 7 or 8 criteria: trade openness, business cycle synchronization, inflation convergence, real exchange rate volatility, and real interest rate cycle symmetry. For Maastricht, 3 of the 4 criteria: inflation convergence, interest rate convergence, and nominal exchange rate variability.

Variable measurements for currency basket anchor and for euro anchor are exactly the same with those for other anchors. However, there are slight adjustments when the monetary anchors involve the euro whose reference country is EMU (recall that in this study the alternative anchors considered are: US dollar, Japanese yen, G3 currency basket, euro, and Chinese yuan). For the trade variable, bilateral trade intensity with EMU is measured vis-à-vis 11 founding member countries of EMU. For the exchange rate variable, the German mark is used as the reference for periods before advent of euro. For the same reason, the German interest rate is used for the time before the marginal lending facility rate for eurozone was put in place. For the business cycle and inflation variables, Germany is set as the reference throughout the periods.

Preliminaries

Prior to the main analysis, there is preliminary analysis in which the features of the data are reviewed descriptively to detect any interesting patterns, particularly any common patterns over the countries across monetary anchors and periods.

5.6 The Data Analysis Process

This section summarizes the procedures involved in producing the results based on the OCA and the Maastricht criteria. A much clearer understanding will be gained subsequently in the respective chapters, Chapters 6 and 7. When interpreting the results, special attention is given to the findings involving the countries associated with fixed exchange rates in practice, namely Hong Kong, Macau, Brunei, and Singapore. This will meet objective nine, the last research objective.

Throughout the analysis, comparisons of findings are done across monetary anchors and periods. Except for the preliminaries, results will also be evaluated over cluster analysis methods. Remember that the labor data are only available for the post-crisis period, hence analyses without and with the labor criterion are performed for the post-crisis period.

The data analysis process begins with preliminary analysis. Then, analyses using HCM, FCM, and MBC methods are performed sequentially using OCA criteria before being repeated with Maastricht criteria. For OCA criteria, clustering analyses are also repeated using weighted criteria.

The process proceeds with assessment of preparedness. For OCA criteria, dollarized and euroized cases play the role of benchmarks whilst for Maastricht criteria, only euroized cases serve as the benchmarks. For OCA dimensions, the preparedness assessment is also carried out with weighted criteria.

Towards the end, comparisons between OCA and Maastricht results are made. The original cluster analysis solutions will also be assessed against the PCA findings. This segment is located at the end of Chapter 7.

Aside from the main analysis above, subsidiary analysis is also carried out using OCA criteria. In the first part, subsets of criteria which are more dominant in molding the classification solutions are determined. In the second part, subsets of criteria corresponding to the best partitions as indicated by the validation measures are identified. This is presented at the end of Chapter 6.

5.7 Chapter Conclusion

This chapter has lined up the procedures involved in achieving the research objectives whose results are expected to answer the respective research questions.

To meet the first objective, findings using different monetary anchors will be compared and contrasted. For the second objective, two sets of criteria are used alternatively. The discussion on the monetary anchors and the criteria has been presented in the previous chapter.

This chapter began with descriptions on the pattern recognition techniques which are needed specifically for objective three. The chapter has also explained how the fourth objective can be attained, that is, by weighting the benefit and cost aspects of monetary unification equally. Next, by inserting certain benchmarks in the analysis, the levels of preparedness for integration can somehow be assessed, meeting objective five. Subsidiary analyses on the OCA criteria will be run to attain objectives six and seven. To accomplish the eighth objective, data sample is sliced into three distinct but successive time periods. Finally, for the ninth objective, special attention is given to cases with prevailing fixed exchange rate arrangements.

The ensuing chapter, Chapter 6 Results using OCA Criteria presents the results using OCA criteria.