Research Article

The Industry Similarity in Input-Output System of China in 1981-1995: Application of Dual Scaling and Fuzzy Clustering

Xue Fu[*]

Nanchang University, Nanchang, China, 200234

[*]**Corresponding author**: Xue Fu, Shanghai Normal University, Handan Roadv 220, Shanghai, P.R.China, Tel: 13970831700, E-mail: fuxueshanghai@yahoo.com.cn

Citation: Xue Fu (2024) The Industry Similarity in Input-Output System of China in 1981-1995: Application of Dual Scaling and Fuzzy Clustering. J Artif Intel Soft Comp Tech 1: 1-23

Abstract

This paper address the evolution of the similarities between industries from 1981 to 1995 in China by applying dual scaling and fuzzy clustering to input-output systems. The dual scaling is applied to set simultaneously in row and column direction to the simple structure of demand and supply by row and column eigenvector as their weights. The distances of weight between rows reflects the similarity of the industries in sales profile between, and that between columns reflects that in purchases profiles. Because of overlap of cluster of industries, network graph analysis and the fuzzy clustering is applied to find the complex similarity between industries. The consistent findings are as follows:

(1) There are increasing overlap or fuzzy in the similarity both in supply and demand with the development of economy in 1981-1995.

(2) Supply industries were rapidly specialized and demand was divergence stable over this period;

(3) Supply industries were rapidly specialized and demand was divergence stable over this period; while heavy industry, such as Machinery and Equipment, became the industry with the most similar to other industries both from supply profiles and demand profiles.

Keywords: Input-Output System; Dual Scaling; Similarity; Network Graph.

Introduction

Leontief input-output model is a kind of powerful tools to classify and interpret structure. Based on this model, a lot of methods were developed [1-3]. The largest engine value and responding engine vector of input-output model reflects the economic system of equilibrium growth rate and structure (Andras Brody, 1997). Because of input-output table is a multivariate categorical data, so the correspondence analysis can be used to deal with the input-output table to get simple structure in multi-dimension. Principle Component Analysis is one way to get the new coordinate system based on eigenvalue. A reassessment of PCA was made in input-output table both in the row cluster (R-mode) and column clusters (Q-mode) by approach of O'Huallachain (1984), the largest engine value in row direction and column direction corresponds the equilibrium rate and price. Dual scaling, presented by Nishisato [4,5], provides a way of mapping out the relative effects of a change in demand or supply in a sector as a whole. When applying it to the input-output table, the eigen value is simultaneously both for row and column, the corresponding eigen vector is the equilibrium structure both demand and supply.

Dual scaling technology is first used by [6] to simultaneously evaluate the important change originating in the demand for inputs (technological condition) and in the supply (market conditions). Taking row and column eigenvector as their weights, the distances of weight between rows reflects the similarity of the industries in sales profile between, and that between columns reflects that in purchases profiles. Similarity between rows and columns can reflect the interrelation and between different industries in sale profiles and purchase and specialization. The clusters were able to roughly describe the similarity in industries. To further describe the overlap between these clusters, we first describe the serious of overlap by using the network graph, and then employ the fuzzy clustering to address the evolution of similarity of Chinese input-output system from 1981 to 1995. Utilization of the fuzzy clustering to investigate the similarities between demand industries or supply industries provides a useful method to describe the complex similarity among industries in input-output system.

If describing the similarity relationship between industries in the input-output system as a complex network, the network graph can be used to clarify the complex structure of similarity. Because of overlap of the complete sub-clusters, the further fuzzy clustering is to discover the more information on the specialization in the sale profiles and divergence in the purchase profiles. Comparing to the [6] two regions comparison analysis, the second feature of my paper is to address the dynamic evolution of the similarities between industries from 1981 to 1995 in China.

Section 2 presents the link from the due scaling, network graph to fuzzy clustering. In section 3, we start from the analysis of overlap among the clusters using network graph, and then apply the fuzzy clustering to Chinese input-output tables in 1981 -1995 to address the similarity in supply industries and demand industries. We will conclude with a summary of results and further research possibilities.

Link from dual scaling, network graph to fuzzy clustering

When regarding the input-output table as contingency table, the weights of demand (column) and supply (row) are obtain simultaneous by application of dual scaling technology. It is similar to principal components analysis (PCA) for categorical data approach; however, the existence of a dual solution, as will be seen later, makes dual scaling different from the well-known PCA and allows for additional insights. The traditional input-output analysis compared one by one the ripple effect of a change in intensity of key industry [7-10] or of key elements only in technical block in an input-output system [2]. According to the dual scaling method, we can obtain simultaneously both of the weight (principle component) of supply industries (row) and demand industries (column). Based on the two dimensions of weights, further analysis on the similarity will be more precise than tradition analysis one by one, and then the investigation of association will combine the Leontief model and Ghosh model at the same time. We focus on the former instead of the latter. The more detailed description of this method refers to Appendix. Applied to an input-output table, the weights will be used to obtain clusters in sales profiles and in purchase profiles; in order to achieve this, we will disregard the primary inputs (last row) and the find demand (last column) from any further analysis.

Clustering [11] is one of the important techniques

in Data Mining. Cluster analysis divides data into groups (clusters) such that similar data objects belong to the same cluster and dissimilar data objects to different clusters. The resulting data partition improves data understanding and reveals its internal structure. We use the clustering methods to group supply industries or demand industries into clusters. The similarity between two row (or column) weights is defined by a distance function.

The row (or column) weight matrices can be used to compute two sets of Euclidian distances in the space of dimension s, and inter-rows (inter-columns) distances and rows-columns distances as follows. The inter-rows distances are found by computing the following r×rEuclidian distances:

$$d_{i,i'} = \sqrt{\sum_{k=1}^{s}(y_{ik} - y_{i'k})^2} \qquad (1.18)$$

The inter-columns distances are found by computing the following $C{\times}C$ Euclidian distances:

$$d_{j,j'} = \sqrt{\sum_{k=1}^{s}(x_{jk} - x_{j'k})^2} \qquad (1.19)$$

Sectors similarities in component graph

The objective of applying dual scaling is to first find similarities or divergences between sales profiles (rows) and between purchases profiles (columns). Applied to an input-output table, the distance of weights can explain the relationship between industries. In order to explain this relationship, we classify the distances of weights between rows and those between columns, and then find the association of distances of weights between rows and columns. The first type of weight relationship can be analyzed by topology and clusters; the second type weight relationship can be analyzed by assignment problem and graph.

Graph theory, so called topology, investigates the mathematical structure which model pairwise relationship between objects from a certain collection. Euler solved the famous problem of the bridges of Königsberg, which is view as the first result of graph theory. This theory has developed during the second half of the 19th century and has boomed since the 1930s. It applies to Computer Science, Operations Research, Game Theory, Decision Theory. Our motivation is explaining relationship between inter-industries by this new theory. We draw a graph consisting of a collection of vertices (referring to industries) and a collection of edge that connect pairs of vertices (referring to the relationship between industries), and then investigate the evolution of topology structure of the graph standing for the similarity development.

We introduce some definitions and other variations in the types of graphs that are considered in our study, more detailed conceptions refer to [12]. Let G be a *graph*, $V(G)$ the set of vertices of G, $E(G)$ the set of edge of G. *A simple graph* is a graph such that there is at most one edge between any two vertices of G. Let $v_i, v_j \in V(G)$, v_j is adjacent of v_i. if $v_i v_j$ is an edge of G. A *path* is a subgraph of G, denoted by $v_0 v_1 v_2 \dots v_k, k \geq 0$, such that $v_i \neq v_j$, $i \neq j$, $i,j = 0,1,2,\dots,k$ and v_i is *adjacent* to v_{i-1} and $v_{i+1}, i = 1,2,\dots,k-1$. *A connected graph* is a graph such that for every two vertices v_i, v_j of it there is at least a path from v_i to v_j. *An induced subgraph*: Let $V' \subseteq V(G)$, An induced by V' subgraph of G such that for every edge e of it the two vertices incidenced with e are in V', denoted by $G[V']$. *A maximum connected subgraph*(i.e. component) of a graph G is a connected subgraph, denoted by , of G such that for $\forall v \in V(G) \smallsetminus V(H)$, $G[V(H) \cup \{v\}]$ is not a connected. A complete graph, or clique, is a simple graph where an edge connects every pair of distinct vertices. The complete graph on n vertices has *n* vertices and *n(n − 1) / 2* edges, and is denoted by K_n. Any subgraph induced by a clique is a complete subgraph. A maximal clique is a clique that is not a subset of any other clique. Let G_i is maximal clique of G, $\bigcup_{i=1}^{k} G_i$ and $\bigcap_{i=1}^{k} G_i$ are respectively the union and the intersection. A simple graph can be express as the union of all maximal cliques.

To assess the similarities or dissimilarities between sale profiles (rows) or purchase profiles (column), we consider that distances below a given threshold are a good indicator of similarity between sectors. The threshold can be fixed arbitrarily; however, in our case we will fix its level at the 1st quartile. Taking sale profiles for example, we choose the elements under threshold from the row weight distance matrix. Responding to a certain element, its two industries are two vertices (v_i and v_j) with an edge connecting them(e_{ij}) standing for the similarities between two industries' sale profiles (row). From all elements under threshold, we draw the responding simple graph, which consist of several connected graph. Some isolated vertices (industries) are not adjacent to other vertices (industries), so they are dissimilar to other industries. The simple graph and those isolated vertices compose a graph G.

From the graphs, we observe the similarities between sale profiles by examining the properties of the graphs.

First, we examine the *degree* d_i of a vertex v_i, (i.e. the number of edge which touch). The more degree the vertex v_i has, the higher similarity the industry i has to other industries. We sort all vertices in decreasing order of degree and get a degree sequence $D(G) = (D_1, D_2, \ldots, D_N)$, then according to the *maximum vertex degrees*, *means of degree*, *variance of degree*, and the number of *isolated vertices*, we infer the industries the most similar to industries, the means and variance of the industries' similarity.

Second, we examine the strength of connectivity in a graph. It can be in particular captured by properties of the maximum cliques in the above simple graph. We compute the number and length of maximum clique, and obtain their intersection part. The strength of connectivity can clearly describe the similarity between the sale profiles. The closest similarities in sale profiles are summarized in table 1. The involved several clusters can be regarded as a set of a labeled maximum cliques with vertex set. All maximum cliques form a simple graph. Their intersection concerns about n(n-1)/2 common edges if it is a clique of order n. Thus the order the intersection is $O(n^2)$ and is critical to the strength of connectivity of the graph. For example, we analyze the strength of connectivity for sale profiles in 1995. In figure 1, we can observe clearly the intersection of
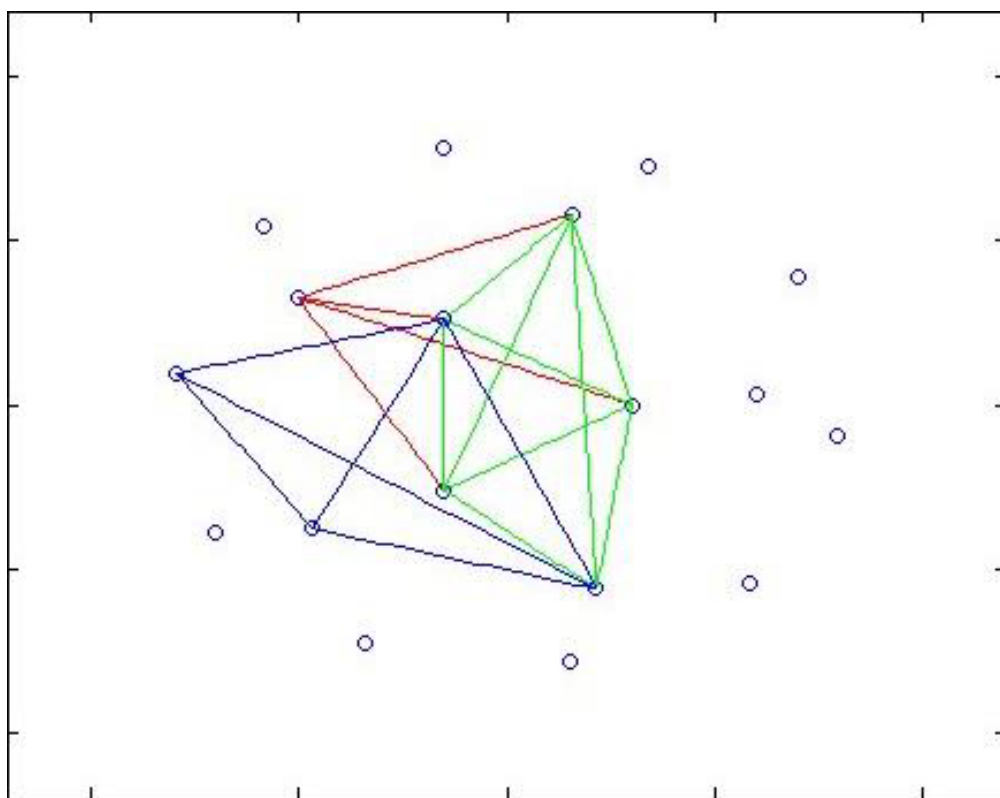

Figure:

Third, we see the evolution of topology structure. Due to the complexity of graph, we compare the properties of topology, in particular the intersection of the maximum cliques. From the Figure 1, we can see the evolution of similarity in sale profiles from 1981 to 1995. The evolution of purchase profiles is shown in Figure 2.
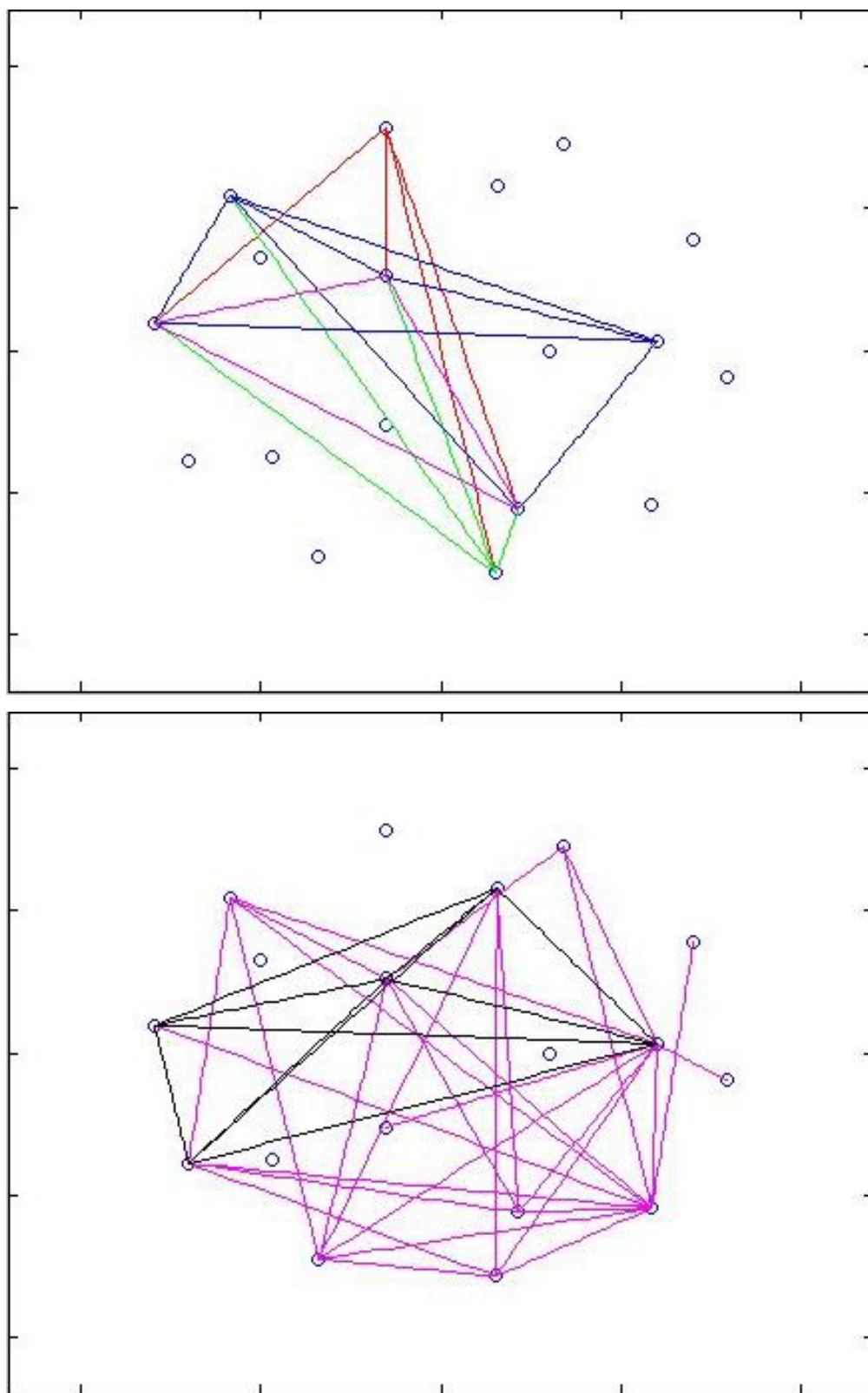

Figure:

those in column weight distance matrix stand for the similarities between demand industries (purchase profiles). Based on the distance tables and the chosen threshold we observe the clusters summarized in Tables 1 and Figure 1 (the initial clusters are provided in Table A.32 in the Appendix).

Table 1: Secttors clustering for China 1981-1995

	Similarities in Sales Profiles	Similarities in Purchase Profiles
1981	{AGR, PCE} {AGR, CHM, MEC, FDI, WVI, BFI, NMP}	{AGR, CHM, WVI, CST, BFI} {AGR, CHM, CST, TRN, BFI}
	{AGR, CHM, MEC, WVI, TRN, BFI, NMP}	{AGR, FDI, WVI}{AGR, MEC, CMI}{AGR, MEC, WVI, CST}
	{AGR, MEC, FDI, WVI, CST, BFI, NMP}	{AGR, MTG, MEC, CST} {AGR, OIL, CHM, TRN}
	{AGR, FDI, WVI, LTR, CST, NMP}	{AGR, CHM, CMI, TRN}{WVI, LTR} {COK, CMI}
1983	{AGR, CHM, MEC, PCE, BFI, NMP}	{AGR, CHM, CMI, CST, TRN} {AGR, CHM, CMI, WVI, BFI}
	{AGR, CHM, MEC, FDI, TRN, BFI, NMP}	{AGR, MEC, CMI} {AGR, COK, CMI, TRN}
	{AGR, FDI, WVI, BFI, NMP}{AGR, MEC, FDI, LTR, CST, NMP}	{AGR, WVI, PCE}{AGR, FDI, WVI} {AGR, OIL, TRN}{AGR, OIL}
1987	{AGR, CHM, MEC, FDI, WVI, BFI, NMP}	{AGR, CHM, MEC, TRN, BFI} {AGR, MEC, CMI, TRN, BFI}
	{AGR, MEC, FDI, WVI, LTR, CST, BFI, NMP}	{AGR, CHM, FDI, WVI, BFI}{AGR, CHM, WVI, TRN, BFI}
	{MEC, WDF, BFI}{TRN, BFI}	{AGR, MEC, CMI, WDF}{AGR, COK, CMI} {WVI, LTR, BFI}
	{OIL, CHM} {CMI, WDF}	{WVI, LTR}{MTG, MEC, CST}{MTG, MEC }{ MEC, WDF, CST}
1990	{AGR, CHM, MEC, FDI, WVI, BFI, NMP}{TRN, BFI} {CMI, WDF}	{AGR, CHM, MEC, WVI, BFI}{AGR, COK, MEC, CMI}
	{AGR, MEC, FDI, WVI, CST, BFI, NMP} {MTG, MEC}	{AGR, CHM, MEC, TRN, BFI}{AGR, CHM, MEC, WDF}{MEC, WDF, CST}
	{AGR, CHM, MEC, WDF, BFI} {AGR, FDI, LTR, CST, BFI, NMP}	{AGR, CHM, FDI, WVI, BFI} {AGR, MEC, CMI, WDF}{WVI, LTR}
1992	{AGR, MEC, FDI, PCE, TRN, BFI, NMP}{AGR, WVI, BFI}	{AGR, MEC, PCE, BFI} {AGR, MEC, PCE, TRN} {AGR, MEC, CMI, BFI}
	{AGR, MEC, FDI, CST, BFI, NMP} {AGR, CHM, MEC, BFI, NMP}	{AGR, MEC, CMI, WDF, CST}{AGR, MEC, CMI, TRN}{AGR, CHM, CMI}
	{AGR, LTR, NMP}{MEC, CMI, BFI} {MEC, CMI, WDF}{OTI, TRN}	{AGR, WVI, PCE}{AGR, OTI, BFI}{AGR, FDI} {MTG, MEC }
1995	{AGR, MEC, FDI, CST, BFI} {AGR, MEC, FDI, PCE} {FDI, LTR}	{AGR, MEC, FDI, BFI} {AGR, COK, MEC, CMI}
	{AGR, MEC, FDI, TRN, BFI} {AGR, MEC, CMI, NMP}{ MEC, WVI}	{AGR, CHM, MEC, FDI}{CHM, MEC, CMI, CST}
	{AGR, MEC, FDI, CST, NMP} { MEC, CMI, WDF, BFI}	{AGR, CHM, MEC, CMI}{ MEC, CMI, WDF, CST}{MEC, PCE}

	{AGR, CHM, MEC, CMI, BFI} {MEC, WDF, FDI, BFI}{MTG, MEC }	{ MEC, CST, BFI}{ MEC, OTI}{ MEC, WVI}	
	{AGR, CHM, MEC, FDI, BFI}{ MEC, WDF, FDI, NMP}	{MTG, MEC, CST}{COK, MEC, CMI, WDF}	
	{AGR, MEC, CMI, TRN, BFI}{ MEC, CMI, WDF, NMP}		

Similarities in fuzzy classification

From Tables 1 and Figure 1 it is clear that there is a serious overlap in the sets of similarities of both the sales and purchases profiles for the Chinese economies. Since some sectors belong to more than one set of similarities, we referred to postpone the use of the term 'clusters' until we introduce the notions of fuzzy sets and fuzzy logic.

Often, similarity studies and cluster analysis are used for aggregation purposes (Blin and Cohen 1977), where each industry has to belong to a unique cluster: this should not obscure the option about the possibility of overlap in clusters. In order to investigate clusters overlaps, we can use fuzzy set theory where set distinctions are not crisp (see Miyamoto 1990; Mirkin 1996).

Zadeh (1965) introduced fuzzy sets and fuzzy logic in a celebrated paper. Zadeh, the founder and the most important contributor to the theory, defines a fuzzy set as being "a class of objects with a continuum of grades of membership." Indeed, in fuzzy set theory, unlike in the 19th century Cantorian set theory, elements belong to a set according to a membership function that takes values in the range [0,1] instead of taking a binary value from {0,1},

making Cantorian sets a particular case of fuzzy sets with memberships of 0 or 1.

Unlike the crisp clustering method, fuzzy clustering is concerned with pattern discoveries between sectors rather than their aggregations, it is a discovery tool rather than a summarizing or simplification tool. The use of fuzzy set theory in social sciences in general and economics more particularly, helps avoiding certainties in conclusions based on data that might be inaccurate because of sample bias or collection errors. Ragin (2000) suggests that the ability of fuzzy sets to explore and express a greater diversity in the data is useful to bring closer data analysis and theoretical models in social sciences, since "fuzzy sets can be carefully tailored to fit theoretical concepts".

Formally, let X be a reference finite and countable space of points, where a generic point is denoted by X. A fuzzy set (or subset) A in X is characterized by a real valued membership function $\mu_A(x)$ that associates with each point in $x \epsilon X$ a value from a real interval usually normalized to [0,1]. For notational clarity and to distinguish between fuzzy sets and crisp sets, it is common to denote a fuzzy set A by $\tilde{A}$ (Kaufmann 1975). The fuzzy subset $\tilde{A}$ of X is a set of ordered pairs $\{(x|\mu_{\tilde{A}}(x))\}, \forall x \in X$ where $\mu_{\tilde{A}}(x)$ is the grade or degree of membership of x in $\tilde{A}$.

$$\mu_A : X \to [0,1]$$

If we denote by $\tilde{A}_k, \forall k = 1, \ldots, K$, all the subsets of the universal set X, then the following properties always hold:

$$\begin{cases} \mu_{\tilde{A}_k}(\mathbf{x}) \in [0,1]; \forall \mathbf{x} \in \mathbf{X}, \forall \mathbf{k} = 1,...,\mathbf{K} \\ \sum_{k=1}^{K} \mu_{\tilde{A}_k}(\mathbf{x}) = 1 \end{cases}$$

One reason for the lack of use of fuzzy logic and fuzzy set theory in social sciences may be attributed to the problematic way to determine the shape of the membership

function, an issue that has not received sufficient treatment in the literature. Indeed, although assuming an ad-hoc shape for the membership function–as it is done when dealing with

probability distribution functions–might be convenient for theoretical studies, it constitutes a serious departure from the data's underlying patterns when used for applied work. An example of ad-hoc defined probability distribution is used in Jackson (1986) to estimate technical coefficients in input-output tables. Kaufman and Rousseeuw (1990, Ch. 4) advance an algorithm, retained in the statistical package S-Plus 2000 (MathSoft 1999), to compute the membership values5.

The iterative algorithm proposed by Kaufman and Rousseeuw (1990) consists in classifying r objects (here sectors) in k clusters based on the observation of s characteristics, variables, or observations of a variable (here the dual scaling solutions). With $d_{i,i'}$ the distance computed in (29) for the supply profiles and (30) for the demand profiles, for each sector i and cluster k there will be a membership μ_{ik} value that solves the following program:

$$\min_{\mu_{ik}} \sum_{v=1}^{k} \frac{\sum\limits_{i,i'} \mu_{iv}^2 \mu_{i'v}^2 d_{i,i'}}{2 \sum\limits_{i'=1}^{r} \mu_{i'v}^2}$$

$$s.t.$$

$$\mu_{iv}^2 \geq 0; \forall i = 1, \ldots, r; \forall v = 1, \ldots, k$$

$$\sum_{v=1}^{k} \mu_{i'v} = 1; \forall i = 1, \ldots, r$$

It allows a maximum number of clusters $k = \frac{n}{2} - 1$, where in our case n is the number of industries (i.e., the size of the technical block). Aside from numerical considerations proper to the algorithm, the maximum number of clusters is set so as no sector represents a unique cluster and that the number of sectors is not equal across clusters. However, the optimal number of clusters has to minimize the objective value in (11). To assess the fuzziness of the resulting cluster Dunn's partition coefficient is computed:

$$F_k = \sum_{i=1}^{r} \sum_{v=1}^{k} \frac{\mu_{i'v}^2}{r}$$

with $F_k \in [\frac{1}{k}, 1]$, such that for entirely fuzzy clustering $\mu_{i'v} = \frac{1}{k}$ and $F_k = \frac{1}{k}$, and for entirely crisp sets $\mu_{i'v} = 0$ or $\mu_{i'v} = 1$ and $F_k = 1$. A normalized Dunn's coefficient taking values from[0,1] is computed by:

$$F_k^= \frac{F_k - \frac{1}{k}}{1 - \frac{1}{k}} = \frac{kF_k - 1}{k - 1}$$

The fuzzy clustering of Chinese industry similarity in 1981-1995

The objective of applying dual scaling is to find similarities or divergences between sales profiles (rows) and between purchases profiles (columns). Applied to an input-output table, the distance of weights can explain the relationship between industries. In order to explain this relationship, we classify the distances of weights between rows and those between columns, and then find the association of distances of weights between rows and columns. The first type of weight relationship can be

analyzed by topology and clusters; the second type weight relationship can be analyzed by assignment problem and graph.

Overlap in clusters

Using the row and column weights solutions, we can compute the Euclidian distances between pairs of row weights and then the Euclidian distances between pairs of column weights as described in the distance matrices. The closest similarities under threshold are summarized into a set of clusters shown chronologically in table 1.

Table 2: Dunn's index of fuzziness

	Sales Profiles			Purchase Profiles		
	Objective	Dunn's coeff.	Dunn's norm.	Objective	Dunn's coeff.	Dunn's norm.
1981	4.8159	0.4650	0.3886	5.2707	0.4593	0.3820
1983	4.9803	0.4580	0.3806	4.8946	0.4608	0.3838
1987	4.9566	0.4652	0.3888	5.6059	0.4614	0.3844
1990	6.0942	0.3956	0.2948	5.5754	0.4619	0.3851
1992	4.8222	0.4568	0.3792	5.1881	0.4610	0.3840
1995	5.2199	0.4571	0.3796	5.8333	0.4598	0.3827

The table 1 shows each set of clusters may have overlap sub-cluster. To observe intuitively, they are shown as the simple graphs consisting 18 nodes (industries) and edges connecting the two industries, the distance of which are under thresholds. The similarity evolution is shown in Figure 1 for the sale profiles and in Figure 2 for purchase profiles. In order to show the overlap in the largest degree, we observe the maximum intersected cluster overlapped by the maximum clusters. From the evolution of clusters and their maximum intersection parts, we find is that production became rapidly specialized and requirements were gradually divergent from 1981 to 1995. As to the sale profiles, the number of maximum cluster increase from 5 in 1981 to 8 in 1995. The most of orders of maximum cluster are 6-7 in 1981, while those in 1995 are about 3. The marketable degree is weak in the early of 1980s, production structure is simplified so that supply industry is similar, and. With the development of market economy, the industries were increasingly specialized, so there were many small groups of industries, which were increasing similar in supply profiles. This is can be prove by searching the maximum intersection clique of the variety of combination of aforementioned maximum cliques. We stand out the intersected cluster with maximum order, and their parent cliques by star. The maximum of intersected clusters is marked by the + line, and the parent cliques are marked with different colors. In intersected maximum clusters, the core of supply industries are consistent over the whole period, they concern about the fundamental industries and heavy industries, such as, Agriculture, Machinery and Equipment, Foodstuff, and Commerce and Catering Trade.
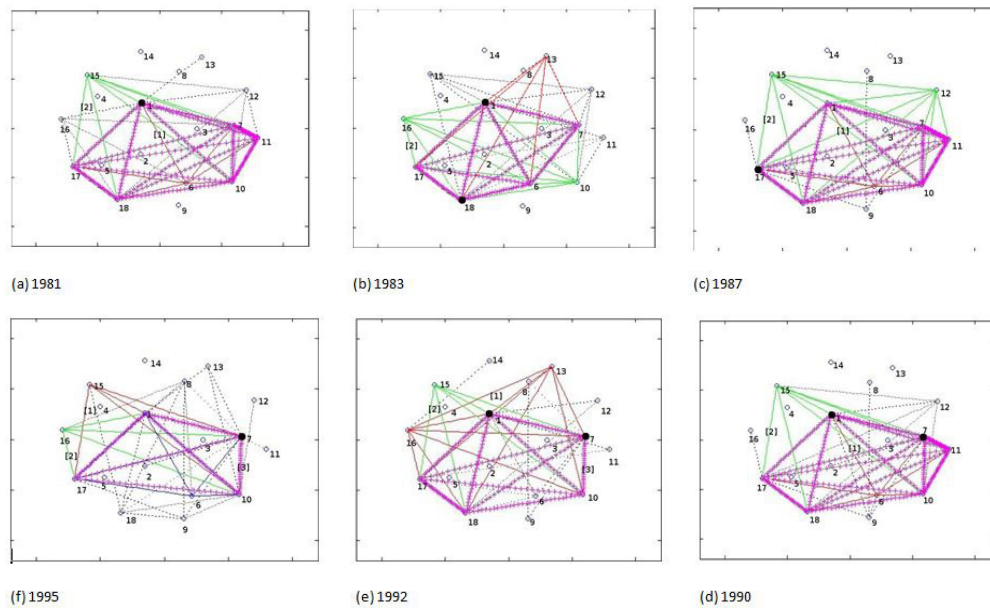
Figure 1: The evolution of similar in sales profiles

As to the purchase profiles, industry similarity is more stable. Although the number of maximum clusters are higher than supply profile in 1981, but it is more stable and less influence by non-market factors. With the development of market economy, most industries similarity increase the stable and were involved the requirement of households, such as industry 1, 6, 8, 15, and 17. Because of divergence, the intersected cliques were less than that of supply profiles. From Figure 2, we find that the core industries in purchase profiles, are the heavy industry 7, agriculture, and commerce became the core of industries with similar demand. In intersected maximum clusters, the core of demand industries is inconsistent over the whole period, they convert from the Agriculture, Transportation, Post and Telecommunications, and Commerce and Catering Trade to the heavy industries, such as Machinery and Equipment, Building Materials and Nonmetal Mineral Products. The reason is the requirement in the early of 1980s are plodding, while heavy industries developed to the important industries and investment demand became the main derive of development of economy.
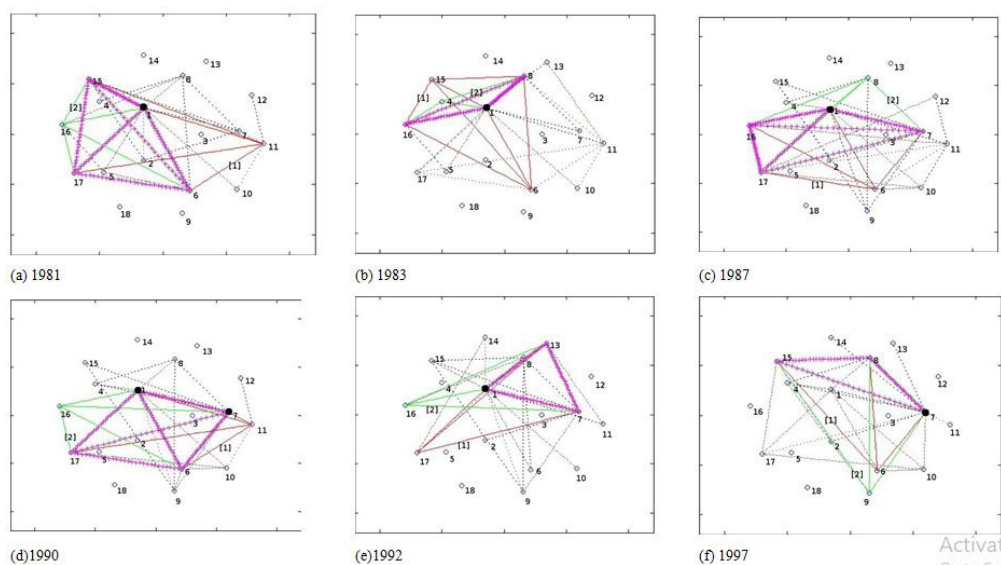
Figure 2: The evolution of similar in purchase profiles

Network graph is promoted to address the degree of overlaps between these clusters. According to graph theory, we draw network graphs of both sale profiles and purchase profiles under threshold of distance among the row weights and column weight. The network graph consists of a set of points called vertices (singular is vertex) connected by lines called edges. Network analysis has emerged as a key technique in modern economics and sociology, and study complex sets of relationships (Stanley and Katherine, 1994). The row and column weight distance connote the similarity between the supply industries and demand industries. From a set of network graphs chorological, we analyze the evolution of Chinese industry similarity in 1981-1995. [14] merely gave the clusters whose distances are below threadholds. As for more detailed industries over long period, it is not enough to clarify the complexity of similarity. It is necessary to address the property of graph, including the statistic index of degree, and marking the maximum cliques and their intersections. These analyses will offer more clear explanation of evolution of similarities among the supply industries and demand industries.

The fuzzy clustering of similarity of Chinese industry similarity in 1981-1995

The overlap between the clusters requires the further fuzzy clustering. We address the similarities or divergences between sectors of supply (rows) or demand (column) by fuzzy clustering [12] analysis the similarity for 10 industries two regions by means of fuzzy clustering because there is overlap in the clusters with the closest distance. We survey fuzzy clustering in 18 detailed industries over long period from 1981 to 1995. Before analyzing the fuzzy clustering, we are required to clarify whether they overlap in the similarity clusters with closest distance. Each cluster means clump, of which the nodes are similar to each other for the sale profiles (or purchase profiles).

Similarities in fuzzy classification

From above analysis, it is clear that there is a serious overlap in the sets of similarities of both the sales and purchases profiles for the Chinese economies.

Applying the above algorithm to the sales and purchases profiles of the China provides the memberships to six fuzzy sets, the maximum number of sets was chosen by

the algorithm. However, lowering the number of sets to five, for example, proved to yield non-optimal values to the objective function in (11). The membership values are provided in Tables A.2- A.3 of the Appendix. One way to compare the overall clustering observed for the China from 1981 to 1995 is to look into Dunn's index in Table 2. In Table 2, Dunn's index suggests that indeed the clusters are lower fuzzy for China (0.29-0.39 for sale and 0.38 for purchase) comparing US (0.24 for sale and 0.21 for purchase) and Canada (0.21 for sale and 0.13 for purchase); further, the table indicates that, as for the sales profile, the degree of fuzziness in the clustering slowly decrease from 0.3886 in the early of 1980s' to 0.3888 at the end of 1980s', and suddenly became greatest in 1990 (0.2948), and then come back quickly to 0.3796 in 1995, and as for purchase profile, the degree of fuzziness in clustering from increase gradually from 0.3820 in 1981 to 0.3827 in 1995 but became lowest in 1990 in contrast. In addition to being a supplemental proof of the inappropriateness of the use of crisp clusters for comparative studies, Dunn's index suggests that the difference in fuzziness implies the difference in specialization and market degree in Chinese industries between supply and demand. Excluding the non-market factor, the supply of Chinese industries became more and more specialized and marketable from the early to the end of 1980's. It suddenly became lowest around 1990 due to non-market factor and then quickly recovered high again. Comparing to the former, the demand of Chinese industries also gradually deepened the specialized and marketable degree and more stable to be influenced by the non-market factor over the same period. As known, China implemented market economy since the end of 1970's or the early of 1980's. The result reflects that before 1995 Chinese market economy was weak but it destined to intensify, in addition, it was an asymmetric between supply and demand. Before 1995, the degree of state-owned was high, thus supply are apt to be affected by non-market factor comparing to demand. Indeed, the existence of overlaps in the clusters is an indicator of the difficulty of drawing a line separating economic sectors that belong to different clusters, or to polarize industries. It implies that supply of industries in China evolved more similarity with each other than the demand of industries 1980-1995.

At the cost of a loss of information, we can force each sector to belong to the set in which it holds the highest

membership and look for the closest neighbor to that set; we

obtain a "forced" classification as shown in Tables 3.

Table 3: Dunn's index of fuzziness

Num.	Sale profile						Purchase profile					
	81	83	87	90	92	95	81	83	87	90	92	95
1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1	1
3	1	2	2	1	1	1	2	1	1	1	1	1
4	2	2	2	2	2	2	2	2	2	2	2	2
5	3	3	3	3	3	2	3	3	3	3	3	2
6	3	3	3	3	3	3	3	3	3	3	3	3
7	4	4	3	3	4	3	3	3	3	3	3	4
8	4	4	4	4	4	4	4	4	4	4	4	4
9	4	5	4	4	4	4	4	4	4	4	4	4
10	5	5	5	5	5	5	5	5	5	5	5	5
11	5	6	5	5	5	6	5	5	5	5	5	6
12	6	6	6	5	6	6	6	6	6	6	6	6
13	6	6	7	6	7	7	6	6	7	7	7	6
14	7	7	7	6	7	7	7	7	7	7	7	7
15	8	8	8	7	7	8	8	8	8	8	8	8
16	8	8	8	7	8	8	8	8	8	8	8	8
17	8	8	8	7	8	8	8	8	8	8	8	8
18	8	8	8	7	8	8	8	8	8	8	8	8

Table 3 shows that the both of sales profiles and purchase profiles for the China are mainly, however not exclusively, represented by clusters A1 (including the agriculture, Minerals and metals and Production and supply of electric power, steaming and Hot water) and A8 (including Construction, Commerce and Catering Trade, Service and Primary input). While the purchases profiles are well represented by clusters A1, and A8. A1 mining and Table 5, for the Canadian economy, shows that the sales profiles are well represented by clusters A1, A2, and A3, while the purchases profiles are mainly represented by clusters A1 and A2. The results of Tables 4 and 5 come in confirmation to the Dunn's index values.

Zadeh (1965), in addition to introducing the notion of fuzziness in set theory, introduced elementary operation on fuzzy sets, such as equality, complementation, inclusion, intersection, and union. Consider the following two fuzzy subsets $\tilde{A}_i$ and $\tilde{A}_j$ of the universal set X, whose membership functions are $\mu_{\tilde{A}_i}(x)$ and $\mu_{\tilde{A}_i}(x)$.

Equality. The subsets $\tilde{A}_i$ and $\tilde{A}_j$ are said to be equal if and only if their membership functions $\mu_{\tilde{A}_i}(x)$ and $\mu_{\tilde{A}_i}(x)$ are equal for all $x \epsilon X$.

$$\mu_{\tilde{A}_i}(x) = \mu_{\tilde{A}_j}(x), \forall x \in X$$

Complementation. The complement of a fuzzy set $\tilde{A}_i$, denoted by $\tilde{A}_i^c$, has a membership function:

$$\mu_{\tilde{A}_i^c}(x) = 1 - \mu_{\tilde{A}_i}(x)$$

However, since fuzzy logic violates the excluded middle principle, the intersection of a fuzzy set and its complement does not produce necessarily the empty set.

$$\mu_{\tilde{A}_i}(x) \ \mu_{\tilde{A}_i}(x) \ x \in X$$

Inclusion or containment. The subset $\tilde{A}_i$ is included or contained in $\tilde{A}_j$, and denoted by $\tilde{A}_i \subset \tilde{A}_j$ if and only if:

$$\mu_{\tilde{A}_i}(x) < \mu_{\tilde{A}_j}(x), \forall x \in X$$

Intersection. The intersection of two fuzzy sets $\tilde{A}_i$ and $\tilde{A}_j$ is a fuzzy set $\tilde{B} = \tilde{A}_i \cap \tilde{A}_j$ whose membership function is defined as:

$$\mu_{\tilde{B}}(x) = \min\{\mu_{\tilde{A}_i}(x), \mu_{\tilde{A}_j}(x)\}; \forall x \in X$$

expression (17) can also be abbreviated by:
$\mu_{\tilde{B}} = \mu_{\tilde{A}_i} \wedge \mu_{\tilde{A}_j}$

Union. The union of two fuzzy sets $\tilde{A}_i$ and $\tilde{A}_j$ is a fuzzy set $\tilde{C} = \tilde{A}_i \cup \tilde{A}_j$ whose membership function is defined as:

$$\mu_{\tilde{C}}(x) = \max\{\mu_{\tilde{A}_i}(x), \mu_{\tilde{A}_j}(x)\}; \forall x \in X$$

expression (18) can also be abbreviated by:
$\mu_{\tilde{B}} = \mu_{\tilde{A}_i} \vee \mu_{\tilde{A}_j}$.

If we refer to Tables A.2 and A.3 for the membership function derived from the algorithm suggested by Kaufman and Rousseeuw (1990), we notice no equality between the fuzzy sets in the sales profiles and the purchases profiles. From Tables A.2 and A.3, it is easy to check for the complementation property, and the absence of inclusion in all sets. Unlike the equality between fuzzy sets, the existence of inclusion raises issues of redundancy of the inner set, that efficient algorithms are expected to avoid producing when dealing with fuzzy cluster analysis.

In fuzzy sets, the intersection and union operation offers attractive features, such as the detection of particular import industries in sales and purchases profiles. In Table A. 4 and A. 5, we provide the number of the intersection and union sets membership functions for the sales and purchases profiles for China from 1981 to 1995. The numerical value of membership is between 0 and 1 and their difference is too small to distinguish. Graphing with a logarithmic scale yields the benefit that shows very small value on with a minimum of digits on the graph. Using logarithm along 2, we give a graphic representation with logarithm of the intersection and union sets of membership functions for the sales and purchases profiles for China from 1981 to 1995 in Fig. 3 and Fig. 4. From these figures, we get the same relationship with the original intersection and union sets of membership. To more clearly observe the evolution of membership features of industries, we show both of Figs. 3 and Figs. 4 as the four figures. In figure (a), two line refer to the values after logarithm transformation in 1981(red), and 1995(cyan). Then that in 1987 (blue) is added in figure (b), and that in 1992(black) is added in figure (c). Figure (d) show all the value including that in 1983 (green) and that in 1990 (magenta). From the four figures, it finds that the values after logarithm transformation from 1985 to 1992 are mostly gradually change between that in 1981 and that in 1995. The only unexpected values in 1990 are strange because of non-marketable reason.

For sales profiles, Figure. 3 reveals that the sectors have a very low membership in the intersection of the fuzzy clusters of the sales profiles and a very high membership in the union of all fuzzy clusters. From Figure 3 (a)-(d), the specialized industries are not as important jointly to all the clusters. This strong contrast in memberships shows that: in 1981, those industries 14 OTI, 4 COK and 12 LTR are not as important jointly to all the clusters as they are important to its own clusters 7, 2, and 6 respectively; in 1995, those industries 4 COK, 12 LTR and 2 MTG are not as important jointly to all the clusters as they are important to its own clusters 2, 6, and 1 respectively (see Fig.3 (a)). Industry 4 COK provides product for special industries more than 3

PWP and OIL although they are energy industries; 12 LTR provides more specialized product more than 11 WVI; with the development of industrialization, 2 MTG became the specialized supply industry, while 14 OTI greatly reduced the degree of specialization comparing to that at the beginning of implement of market economy. The heavy industries, i.e. 6 CHM, 7 MEC, 8 CMI, energy industries, i.e. 3 PWR and 5 OIL, construction (15 CST) and services, i.e. 16 TRN, 17 BFI, and 18 NMP, are more important jointly to all the clusters than the light industries, such as, 10 FDI and 12 LTR. Therefore, the supply of those industries may involve in the industries of a lot of industries. For example, providing cost of energy in industry 5 OIL increase, since it belongs to more clusters; its sales will influence more industries.
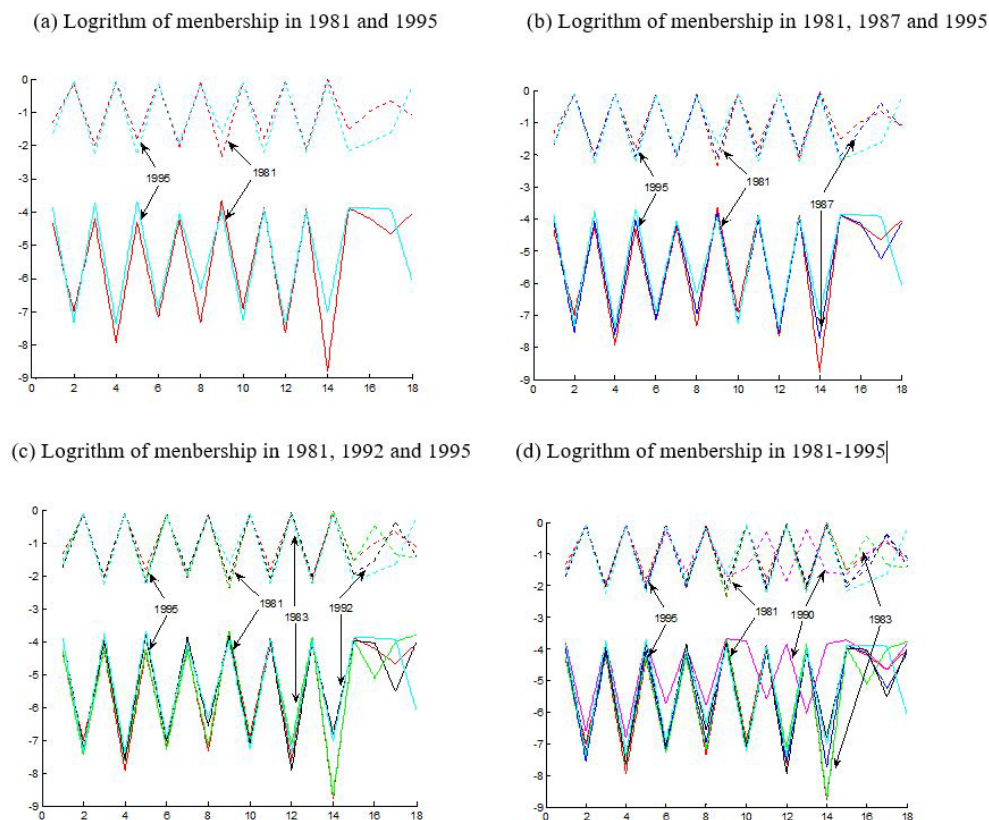
Figure 3: Logrithm of Intersection and Union operation on China fuzzy clusters 1981-1995 for sale profiles

We know the important industries if we compare the changes in the relationship between intersection and union of membership from1981 to 1995. For most of industries, intersection of membership became larger and union of membership became less for most industries. It means these industries supply their products to other industries in more similar structure because they are closely

bound up with each other more and more. However, the similarity of sale profiles is opposite for industry 2 MTG and 10 FDI. The two industries generally provide more particular products to special industries. In contrast, the energy industries, i.e. 3 PWR, 5 OIL have the largest increase in intersection and decrease in union simultaneously. It means that the two energy industry became important because the

degree of similarity to all other industries is increased mostly from 1981 to 1985.

From the figures (b)-(d), it finds that the evolution of similarity increasingly speeds up. For most industries, the line of 1987 was going more near to that in 1992 than that in 1981 (see fig(b)); and the closeness of line between 1983 and 1981 is small than that between 1995 and 1992 (see fig(c)). It is more appearance for industry 14 OTI and 8 CMI. It means the most sale profiles accelerate the speed of similarity for these industries. From fig. (d), the state is instable state in 1990 due to non-market factor, services are in particular instable in similarity, such as 16 TRN, 17 BFI and 18 NMP.

For purchase profiles in Fig. 4, the industries with divergence demand have a very low membership in the intersection of the fuzzy clusters of the sales profiles and a very high membership in the union of all fuzzy clusters. This strong contrast in memberships shows that: in 1981, those industries 14 OTI, 4 COK and 8 CMI are not as important

jointly to all the clusters as they are important to its own clusters 7, 2, and 4 respectively; in 1995, those industries 14 OTI, 4 COK, and 2 MTG are not as important jointly to all the clusters as they are important to its own clusters 7, 2, and 1 respectively (see Fig.4 (a)). As upstream industries, raw mineral resources 4 COK and 2 MTG demanded for special products; as downstream industries, industries 14 OTI and 12 LTR demanded distinguished comparing with the most basic industries. With the improvement of living standard and exploration of new materials, the purchase became increasingly divergent especially for raw material industries, i.e. 2 MTG, 4 OIL and 6 CHM, and for the consumption industries, i.e. 14 OTI, 12 LTR and 10 FDI. There are more important jointly to all the clusters in purchase profiles for the primary industry (1 AGR) and many basic industries (i.e. 6 CHM, 7 MEC), principle energy industries (i.e. 3 PWR and 5 OIL), construction (15 CST) and services (i.e. 16 TRN, 17 BFI). Therefore, the demand of those industries may be more similar and involve in a lot of industries.
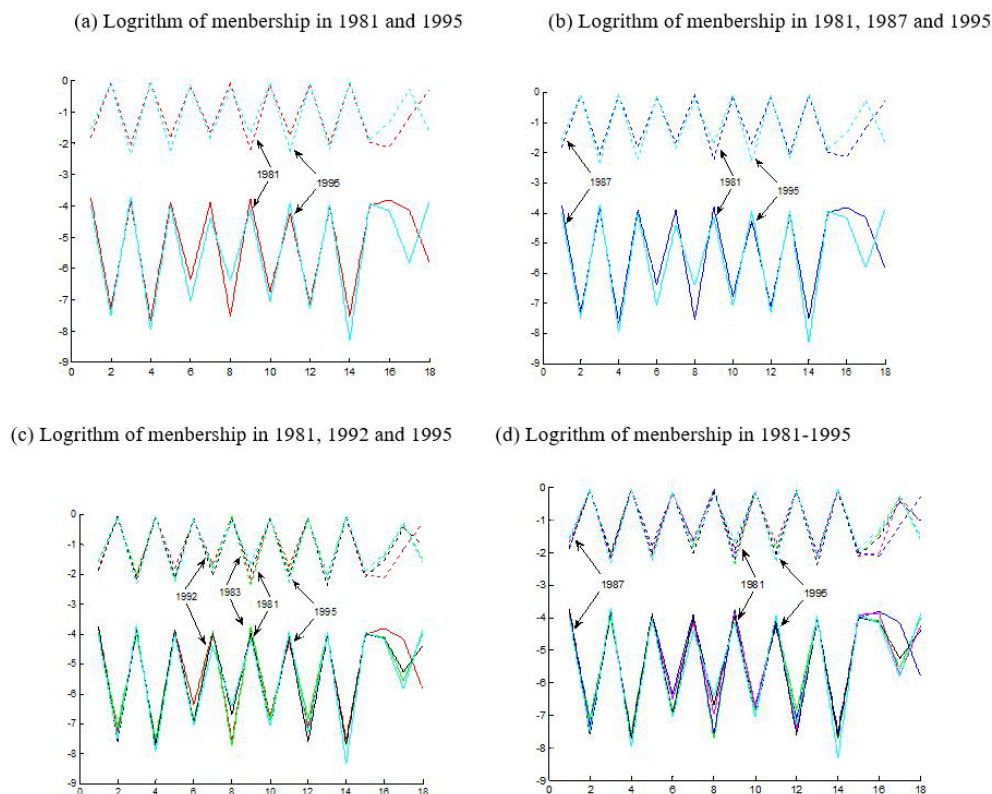
(a) Logrithm of menbership in 1981 and 1995

(b) Logrithm of menbership in 1981, 1987 and 1995

(c) Logrithm of menbership in 1981, 1992 and 1995

(d) Logrithm of menbership in 1981-1995

Figure 4: Logrithm of Intersection and Union operation on China fuzzy clusters 1981-1995 for purchase profiles

We know the demand change of these industries if we compare the changes in the relationship between intersection and union of membership from1981 to 1995.

The tendency of purchase profiles is contrary to those of sale profiles. For most of industries, intersection of membership became small and union of membership became large. It

means these industries divergence their demand for other industries with the advance of society. The demand divergence changed significantly especially for modern material industries, namely 6 CHM, and the special manufactures, 14 OTI. However, the purchase profiles are more important to other clusters for industry 8 CMI, 11 TEX and 3 PWR. The three industries generally developed greatly and demand for the general products from the most of industries.

Conclusion

The paper addresses the industry similarity by fuzzy logic approach for the sale profiles and purchase profiles from 1981 to 1995 in China.

On basis of dual scaling, the priority in both of supply and demand can be calculated according to the 1981-1995 Chinese input output tables. Then we calculate the distance of weights of row or column between different industries. The closest distance under threshold shows the clusters of similarity of sale profiles and purchase profiles. The fuzzy clustering is employed in finding similarity and dissimilarity as well.

Both of them have given the consistent conclusion:

1. There is more overlap or fuzzy with the development of economy in 1981-1995.

2. Supply industries were rapidly specialized and demand was divergence stable over this period; while heavy industry, such as Machinery and Equipment, became the industry with the most similar to other industries both from supply profiles and demand profiles.

3. The purchase profiles evolved more stable than supply profiles when facing non-market factor.

References

1. Sonis M, Hewings GJD (1993) Hierarchies of regional sub-structures and their multipliers within input- output systems: Miyazawa revisited, Hitotsubashi Journal of Economics 34: 33-44.

2. Sonis M, Hewings GJD (1992) Coefficient change in input-output models: Theory and applications. Economic Systems Research 4: 143-57

3. Sonis M, Hewings GJD, Guo J (2000) A new image of classical key sector analysis: Minimum information decomposition of the Leontief inverse. Economic Systems Research 12: 401-23 sale profiles in 1995.

4. Nishisato S (1980) Analysis of Categorical Data: Dual Scaling and its Application (Toronto, University of Toronto Press).

5. Nishisato S (1994) Elements of Dual Scaling: An Introduction to Practical Data Analysis (Mahwah, NJ, Lawrence Erlbaum Associates).

6. Dridi C, GJD Hewings (2003) Sectors associations and similarities in input-output systems: An application of dual scaling and fuzzy logic to Canada and the United States, The Annals Regional Science 37: 629-56.

7. Rasmussen PN (1956) Studies in Inter-Sectoral Relations (Amsterdam, North-Holland).

8. Cella G (1984) The input- output measurement of interindustry linkages, Oxford Bulletin of Economics and Statistics 46: 73-84.

9. Cella G (1986) The input- output measurement of interindustry linkages: a reply, Oxford Bulletin of Economics and Statistics, 48: 379-84.

10. Clements BJ (1990) On the decomposition and normalization of interindustry linkages, Economics Letters 33: 337-40.

11. Tryon RC (1939). *Cluster Analysis*. Edwards Brothers.

12. JA Bondy, USR Murty (1976) Graph Theory with Applications. Elsevier North-Holland.

13. Wasserman, Stanley, Faust, Katherine (1994) Social Networks Analysis: Methods and Applications. Cambridge: Cambridge University Press.

14. Dridi C, GJD Hewings (2002) An investigation of industry associations, association loops and economic complexity: application to Canada and the United States, Economic Systems Research 14: 275-92.

15. Aldenderfer MS, Blashfield RK (1984) *Cluster Analysis*, Sage Publications, Newbury Park, Cal.

16. Hirschman AO (1958) The Strategy of Economic Development (New Haven, CT, Yale University Press).

Supplementary Tables

Table A.1. Input-output aggregation for the Chinese economies

Number	Aggregate	Description
1	AGR	Agriculture
2	MTG	Minerals and metals
3	PWR	Production and Supply of Electric Power, Steam and Hot Water
4	COK	Coking Refining
5	OIL	Gas and Petroleum
6	CHM	Chemical Industry
7	MEC	Machinery and Equipment
8	CMI	Building Materials and Nonmetal Mineral Products
9	WDF	Wood and furniture
10	FDI	Foodstuff
11	WVI	Textile
12	LTR	Sewing, Leather and Furs Products
13	PCE	Paper
14	OTI	Other Manufacturing
15	CST	Construction
16	TRN	Transportation, Post and Telecommunications
17	BFI	Commerce and Catering Trade
18	NMP	Other nonmaterial production
19	OIO	Primary input or final demand

Table A.2. Membership coefficients of sale profile from 1981-1995

		AGR	MTG	PWR	COK	OIL	CHM	MEC	CMI	WDF	FDI	WVI	LTR	PCE	OTI	CST	TRN	BFI	NMP
1981	1	0.399	0.901	0.249	0.011	0.118	0.015	0.096	0.010	0.079	0.008	0.068	0.005	0.067	0.002	0.067	0.055	0.040	0.060
	2	0.127	0.022	0.205	0.946	0.191	0.020	0.120	0.013	0.098	0.010	0.079	0.006	0.073	0.003	0.072	0.057	0.042	0.064
	3	0.122	0.020	0.151	0.013	0.285	0.905	0.237	0.018	0.122	0.013	0.100	0.007	0.085	0.003	0.082	0.063	0.047	0.070
	4	0.100	0.016	0.121	0.010	0.136	0.022	0.241	0.922	0.197	0.016	0.123	0.008	0.097	0.003	0.088	0.067	0.050	0.075
	5	0.075	0.012	0.084	0.007	0.088	0.013	0.107	0.014	0.195	0.914	0.286	0.014	0.141	0.004	0.103	0.077	0.059	0.084
	6	0.061	0.010	0.066	0.005	0.066	0.009	0.075	0.009	0.133	0.019	0.176	0.943	0.231	0.005	0.112	0.083	0.062	0.086
	7	0.049	0.008	0.053	0.004	0.050	0.007	0.054	0.006	0.082	0.009	0.074	0.009	0.166	0.975	0.129	0.094	0.067	0.095
	8	0.068	0.011	0.071	0.005	0.066	0.009	0.070	0.008	0.093	0.011	0.093	0.009	0.141	0.006	0.348	0.503	0.634	0.466
1983	1	0.352	0.925	0.202	0.014	0.114	0.013	0.088	0.011	0.077	0.008	0.064	0.007	0.069	0.002	0.065	0.029	0.063	0.072
	2	0.147	0.018	0.237	0.930	0.212	0.019	0.116	0.015	0.096	0.009	0.076	0.008	0.078	0.003	0.072	0.032	0.069	0.079
	3	0.121	0.014	0.143	0.016	0.235	0.908	0.237	0.023	0.126	0.012	0.094	0.010	0.091	0.003	0.081	0.034	0.077	0.085

Year	#																		
	4	0.112	0.013	0.129	0.013	0.149	0.023	0.249	0.909	0.183	0.015	0.117	0.012	0.105	0.003	0.089	0.037	0.083	0.092
	5	0.078	0.009	0.086	0.008	0.092	0.012	0.108	0.015	0.200	0.916	0.238	0.021	0.151	0.004	0.101	0.042	0.094	0.099
	6	0.068	0.008	0.075	0.007	0.076	0.010	0.081	0.011	0.144	0.020	0.244	0.917	0.214	0.006	0.119	0.048	0.105	0.106
	7	0.052	0.006	0.055	0.005	0.053	0.007	0.053	0.007	0.082	0.009	0.076	0.013	0.147	0.973	0.134	0.050	0.112	0.100
	8	0.071	0.008	0.074	0.007	0.070	0.009	0.068	0.009	0.092	0.010	0.092	0.013	0.145	0.006	0.340	0.727	0.397	0.367
1987	1	0.311	0.938	0.214	0.014	0.111	0.012	0.090	0.010	0.071	0.007	0.061	0.005	0.062	0.005	0.068	0.057	0.026	0.057
	2	0.163	0.016	0.245	0.935	0.199	0.018	0.119	0.014	0.091	0.008	0.074	0.006	0.070	0.005	0.075	0.063	0.029	0.063
	3	0.125	0.011	0.136	0.015	0.240	0.913	0.242	0.021	0.136	0.011	0.095	0.008	0.082	0.006	0.086	0.069	0.033	0.069
	4	0.108	0.010	0.112	0.011	0.137	0.020	0.237	0.910	0.240	0.014	0.123	0.010	0.094	0.007	0.095	0.075	0.036	0.075
	5	0.085	0.008	0.087	0.008	0.100	0.012	0.105	0.016	0.192	0.925	0.244	0.016	0.120	0.008	0.104	0.084	0.039	0.083
	6	0.074	0.007	0.074	0.007	0.081	0.010	0.083	0.012	0.123	0.018	0.233	0.936	0.190	0.011	0.125	0.095	0.042	0.092
	7	0.061	0.005	0.060	0.005	0.061	0.007	0.058	0.008	0.071	0.008	0.083	0.011	0.257	0.946	0.207	0.123	0.047	0.108
	8	0.075	0.007	0.072	0.006	0.072	0.008	0.067	0.009	0.077	0.009	0.085	0.009	0.124	0.011	0.241	0.435	0.747	0.454
1990	1	0.349	0.902	0.260	0.020	0.126	0.028	0.097	0.020	0.079	0.074	0.021	0.070	0.015	0.069	0.076	0.059	0.040	0.069
	2	0.159	0.026	0.226	0.910	0.219	0.041	0.129	0.027	0.102	0.089	0.025	0.080	0.017	0.077	0.083	0.065	0.044	0.076
	3	0.142	0.021	0.157	0.023	0.259	0.821	0.336	0.040	0.161	0.118	0.034	0.099	0.021	0.089	0.096	0.074	0.051	0.087
	4	0.109	0.016	0.116	0.017	0.140	0.044	0.204	0.844	0.302	0.153	0.043	0.119	0.024	0.103	0.104	0.079	0.054	0.093
	5	0.092	0.013	0.095	0.012	0.106	0.028	0.104	0.032	0.185	0.358	0.808	0.271	0.038	0.144	0.132	0.099	0.068	0.113
	6	0.071	0.010	0.072	0.009	0.074	0.019	0.066	0.019	0.091	0.113	0.040	0.235	0.851	0.332	0.195	0.122	0.080	0.131
	7	0.077	0.011	0.075	0.009	0.075	0.019	0.065	0.018	0.081	0.094	0.030	0.126	0.034	0.187	0.315	0.502	0.662	0.430
1992	1	0.301	0.930	0.245	0.011	0.113	0.011	0.088	0.012	0.071	0.007	0.062	0.004	0.062	0.009	0.063	0.061	0.022	0.062
	2	0.150	0.017	0.211	0.944	0.190	0.015	0.108	0.015	0.088	0.009	0.073	0.005	0.070	0.010	0.069	0.067	0.024	0.067
	3	0.126	0.013	0.136	0.012	0.224	0.919	0.212	0.025	0.131	0.011	0.095	0.006	0.085	0.012	0.079	0.076	0.027	0.077
	4	0.115	0.011	0.116	0.010	0.149	0.021	0.277	0.896	0.233	0.015	0.122	0.007	0.099	0.013	0.089	0.084	0.030	0.084
	5	0.090	0.009	0.087	0.007	0.103	0.011	0.107	0.019	0.196	0.924	0.234	0.011	0.128	0.016	0.099	0.094	0.033	0.092
	6	0.067	0.007	0.065	0.005	0.072	0.008	0.069	0.011	0.107	0.014	0.200	0.951	0.165	0.019	0.115	0.099	0.036	0.095
	7	0.075	0.007	0.070	0.005	0.075	0.008	0.070	0.011	0.091	0.011	0.117	0.010	0.248	0.897	0.259	0.173	0.058	0.144
	8	0.076	0.007	0.071	0.005	0.074	0.008	0.069	0.011	0.083	0.009	0.097	0.008	0.144	0.025	0.227	0.347	0.770	0.380
1995	1	0.317	0.936	0.212	0.014	0.112	0.011	0.081	0.014	0.064	0.007	0.066	0.006	0.064	0.008	0.068	0.067	0.066	0.015
	2	0.160	0.016	0.211	0.934	0.210	0.016	0.116	0.020	0.087	0.008	0.078	0.007	0.073	0.009	0.076	0.074	0.073	0.017
	3	0.116	0.011	0.132	0.014	0.183	0.910	0.264	0.031	0.135	0.012	0.100	0.008	0.086	0.010	0.087	0.084	0.082	0.019
	4	0.108	0.010	0.121	0.012	0.149	0.023	0.245	0.873	0.321	0.015	0.124	0.010	0.100	0.012	0.096	0.092	0.089	0.021
	5	0.087	0.008	0.096	0.008	0.109	0.014	0.104	0.023	0.165	0.929	0.204	0.013	0.128	0.014	0.107	0.104	0.100	0.022
	6	0.069	0.006	0.076	0.006	0.081	0.009	0.066	0.014	0.087	0.012	0.213	0.932	0.204	0.019	0.129	0.126	0.113	0.023
	7	0.068	0.006	0.074	0.006	0.077	0.008	0.060	0.012	0.071	0.009	0.118	0.014	0.214	0.910	0.217	0.186	0.156	0.032
	8	0.075	0.007	0.078	0.006	0.079	0.009	0.064	0.013	0.070	0.009	0.098	0.010	0.131	0.019	0.221	0.266	0.322	0.851

Table A.3. Membership coefficients of Purchase profile from 1981-1995

		AGR	MTG	PWR	COK	OIL	CHM	MEC	CMI	WDF	FDI	WVI	LTR	PCE	OTI	CST	TRN	BFI	NMP
1981	1	0.293	0.934	0.231	0.016	0.117	0.016	0.092	0.006	0.079	0.010	0.058	0.008	0.063	0.004	0.065	0.062	0.018	0.066
	2	0.156	0.016	0.239	0.923	0.237	0.024	0.123	0.008	0.094	0.012	0.069	0.009	0.071	0.004	0.074	0.068	0.019	0.075
	3	0.148	0.014	0.156	0.020	0.253	0.899	0.275	0.011	0.115	0.015	0.087	0.011	0.081	0.005	0.086	0.076	0.022	0.086
	4	0.095	0.009	0.093	0.011	0.111	0.018	0.178	0.947	0.183	0.017	0.094	0.012	0.092	0.006	0.092	0.078	0.022	0.088
	5	0.087	0.008	0.080	0.009	0.085	0.013	0.107	0.009	0.170	0.887	0.313	0.027	0.150	0.007	0.144	0.103	0.028	0.127
	6	0.075	0.007	0.070	0.008	0.071	0.011	0.086	0.007	0.150	0.030	0.205	0.904	0.234	0.009	0.160	0.116	0.030	0.129
	7	0.064	0.006	0.059	0.006	0.056	0.008	0.063	0.005	0.105	0.013	0.081	0.015	0.173	0.955	0.154	0.149	0.032	0.122
	8	0.081	0.007	0.072	0.008	0.070	0.010	0.077	0.006	0.105	0.016	0.094	0.014	0.138	0.010	0.226	0.349	0.830	0.308
1983	1	0.332	0.924	0.251	0.012	0.118	0.014	0.092	0.005	0.074	0.008	0.057	0.009	0.063	0.005	0.061	0.058	0.021	0.062
	2	0.141	0.018	0.224	0.942	0.209	0.018	0.120	0.006	0.090	0.010	0.067	0.010	0.069	0.005	0.068	0.062	0.023	0.068
	3	0.131	0.015	0.142	0.013	0.229	0.911	0.286	0.009	0.115	0.013	0.083	0.012	0.081	0.006	0.080	0.071	0.026	0.080
	4	0.083	0.009	0.086	0.008	0.114	0.015	0.151	0.955	0.195	0.014	0.093	0.014	0.084	0.007	0.086	0.072	0.025	0.080
	5	0.089	0.010	0.086	0.007	0.100	0.014	0.116	0.008	0.175	0.909	0.265	0.026	0.145	0.008	0.126	0.099	0.035	0.113
	6	0.078	0.008	0.075	0.006	0.084	0.011	0.090	0.007	0.144	0.022	0.238	0.894	0.241	0.012	0.158	0.119	0.039	0.122
	7	0.066	0.007	0.062	0.005	0.067	0.008	0.067	0.005	0.102	0.011	0.095	0.018	0.170	0.945	0.190	0.143	0.042	0.122
	8	0.080	0.009	0.075	0.006	0.079	0.010	0.078	0.005	0.104	0.013	0.101	0.017	0.147	0.012	0.231	0.375	0.789	0.354
1987	1	0.281	0.938	0.240	0.010	0.111	0.018	0.090	0.005	0.072	0.009	0.053	0.007	0.060	0.005	0.062	0.071	0.056	0.018
	2	0.142	0.014	0.206	0.949	0.178	0.025	0.112	0.006	0.088	0.011	0.062	0.008	0.069	0.006	0.070	0.078	0.061	0.020
	3	0.144	0.012	0.152	0.012	0.284	0.875	0.315	0.010	0.146	0.015	0.084	0.011	0.083	0.007	0.086	0.089	0.075	0.024
	4	0.089	0.008	0.087	0.007	0.102	0.021	0.137	0.950	0.215	0.015	0.088	0.012	0.094	0.008	0.089	0.096	0.069	0.023
	5	0.100	0.008	0.092	0.007	0.101	0.020	0.114	0.009	0.176	0.896	0.302	0.024	0.127	0.009	0.131	0.116	0.095	0.033
	6	0.086	0.007	0.080	0.006	0.083	0.016	0.090	0.008	0.132	0.027	0.238	0.910	0.206	0.013	0.152	0.137	0.099	0.034
	7	0.074	0.006	0.068	0.005	0.067	0.012	0.067	0.006	0.082	0.012	0.085	0.015	0.242	0.940	0.159	0.183	0.102	0.034
	8	0.084	0.007	0.076	0.005	0.075	0.014	0.075	0.006	0.089	0.014	0.089	0.014	0.121	0.012	0.249	0.229	0.443	0.814
1990	1	0.306	0.940	0.217	0.010	0.109	0.017	0.083	0.008	0.068	0.009	0.058	0.006	0.060	0.005	0.065	0.069	0.019	0.052
	2	0.138	0.013	0.190	0.946	0.201	0.026	0.112	0.010	0.089	0.012	0.071	0.007	0.070	0.005	0.075	0.077	0.021	0.059
	3	0.133	0.011	0.146	0.013	0.252	0.872	0.322	0.017	0.152	0.017	0.097	0.009	0.086	0.006	0.090	0.088	0.025	0.070
	4	0.100	0.009	0.110	0.008	0.123	0.027	0.172	0.923	0.240	0.019	0.110	0.010	0.097	0.007	0.098	0.095	0.025	0.072
	5	0.096	0.008	0.100	0.007	0.102	0.020	0.111	0.014	0.191	0.896	0.301	0.016	0.128	0.008	0.127	0.111	0.030	0.089
	6	0.076	0.006	0.081	0.006	0.076	0.014	0.074	0.011	0.108	0.021	0.183	0.930	0.209	0.012	0.147	0.135	0.031	0.090
	7	0.067	0.006	0.072	0.005	0.061	0.011	0.057	0.008	0.071	0.012	0.088	0.012	0.228	0.945	0.167	0.181	0.033	0.095
	8	0.082	0.007	0.083	0.006	0.075	0.014	0.069	0.009	0.081	0.014	0.093	0.010	0.122	0.011	0.230	0.245	0.816	0.471
1992	1	0.272	0.948	0.223	0.010	0.107	0.011	0.080	0.010	0.064	0.007	0.059	0.005	0.060	0.005	0.062	0.057	0.026	0.048
	2	0.147	0.012	0.205	0.948	0.193	0.017	0.104	0.013	0.081	0.009	0.071	0.006	0.071	0.005	0.071	0.064	0.029	0.054
	3	0.135	0.010	0.148	0.012	0.234	0.912	0.242	0.021	0.124	0.013	0.093	0.008	0.089	0.006	0.085	0.076	0.035	0.065
	4	0.108	0.008	0.107	0.008	0.135	0.020	0.236	0.905	0.268	0.016	0.117	0.009	0.103	0.007	0.097	0.085	0.037	0.070
	5	0.097	0.007	0.092	0.007	0.103	0.013	0.117	0.018	0.182	0.917	0.239	0.014	0.142	0.009	0.123	0.100	0.043	0.085
	6	0.076	0.005	0.072	0.005	0.075	0.009	0.075	0.012	0.108	0.016	0.204	0.935	0.190	0.011	0.146	0.112	0.042	0.090

	7	0.074	0.005	0.069	0.005	0.068	0.008	0.064	0.010	0.080	0.010	0.106	0.012	0.191	0.943	0.184	0.163	0.050	0.099
	8	0.092	0.006	0.084	0.006	0.085	0.010	0.082	0.012	0.093	0.012	0.112	0.011	0.153	0.013	0.231	0.342	0.737	0.490
1995	1	0.339	0.940	0.198	0.008	0.094	0.011	0.066	0.015	0.060	0.007	0.065	0.006	0.064	0.003	0.065	0.055	0.018	0.068
	2	0.122	0.012	0.171	0.951	0.212	0.016	0.091	0.021	0.079	0.010	0.080	0.008	0.078	0.003	0.075	0.065	0.020	0.077
	3	0.116	0.010	0.135	0.010	0.192	0.899	0.266	0.042	0.143	0.015	0.109	0.010	0.096	0.004	0.090	0.076	0.024	0.091
	4	0.112	0.010	0.131	0.009	0.162	0.029	0.278	0.848	0.313	0.019	0.137	0.012	0.112	0.005	0.100	0.084	0.026	0.098
	5	0.096	0.008	0.108	0.007	0.113	0.017	0.120	0.030	0.183	0.914	0.197	0.016	0.137	0.005	0.113	0.092	0.028	0.106
	6	0.075	0.007	0.091	0.005	0.085	0.011	0.070	0.018	0.095	0.015	0.207	0.923	0.219	0.008	0.136	0.108	0.031	0.117
	7	0.062	0.006	0.075	0.004	0.062	0.007	0.047	0.012	0.057	0.008	0.101	0.013	0.151	0.964	0.155	0.122	0.032	0.119
	8	0.079	0.007	0.091	0.005	0.079	0.010	0.062	0.015	0.071	0.010	0.105	0.012	0.142	0.007	0.266	0.398	0.820	0.325

Table A.4. Intersection operation on Chinese Fuzzy clusters of sale profiles 1981-1995

	intersection						union					
	1981	1983	1987	1990	1992	1995	1981	1983	1987	1990	1992	1995
AGR	0.0493	0.0516	0.0609	0.0713	0.0673	0.0679	0.3986	0.3517	0.3110	0.3486	0.3010	0.3172
MTG	0.0078	0.0058	0.0054	0.0104	0.0066	0.0062	0.9013	0.9251	0.9375	0.9022	0.9296	0.9364
PWR	0.0531	0.0553	0.0596	0.0715	0.0645	0.0744	0.2492	0.2367	0.2452	0.2598	0.2452	0.2116
COK	0.0041	0.0050	0.0051	0.0089	0.0051	0.0060	0.9455	0.9295	0.9345	0.9103	0.9439	0.9341
OIL	0.0502	0.0531	0.0611	0.0742	0.0718	0.0773	0.2850	0.2347	0.2401	0.2590	0.2236	0.2096
CHM	0.0069	0.0065	0.0071	0.0188	0.0077	0.0083	0.9052	0.9079	0.9127	0.8208	0.9193	0.9103
MEC	0.0538	0.0527	0.0584	0.0651	0.0687	0.0600	0.2407	0.2489	0.2419	0.3360	0.2769	0.2638
CMI	0.0062	0.0069	0.0081	0.0181	0.0105	0.0124	0.9223	0.9091	0.9099	0.8442	0.8961	0.8734
WDF	0.0790	0.0774	0.0711	0.0787	0.0708	0.0644	0.1972	0.1998	0.2398	0.3022	0.2334	0.3213
FDI	0.0084	0.0078	0.0068	0.0738	0.0073	0.0066	0.9138	0.9164	0.9251	0.3582	0.9240	0.9291
WVI	0.0684	0.0635	0.0611	0.0208	0.0621	0.0661	0.2864	0.2437	0.2443	0.8080	0.2336	0.2128
LTR	0.0050	0.0071	0.0053	0.0695	0.0041	0.0059	0.9428	0.9169	0.9356	0.2713	0.9506	0.9316
PCE	0.0669	0.0689	0.0618	0.0152	0.0622	0.0640	0.2314	0.2135	0.2574	0.8511	0.2475	0.2143
OTI	0.0023	0.0024	0.0047	0.0692	0.0090	0.0077	0.9747	0.9734	0.9463	0.3320	0.8966	0.9101
CST	0.0674	0.0647	0.0677	0.0758	0.0632	0.0680	0.3475	0.3400	0.2411	0.3153	0.2593	0.2208
TRN	0.0546	0.0290	0.0567	0.0589	0.0611	0.0673	0.5028	0.7274	0.4349	0.5016	0.3468	0.2662
BFI	0.0395	0.0628	0.0263	0.0399	0.0218	0.0659	0.6337	0.3965	0.7473	0.6624	0.7703	0.3215
NMP	0.0601	0.0721	0.0571	0.0694	0.0616	0.0149	0.4662	0.3674	0.4539	0.4299	0.3801	0.8512

Table A.5. Intersection and union operation on Chinese Fuzzy clusters of purchase profiles 1981-1995

	intersection						union					
	1981	1983	1987	1990	1992	1995	1981	1983	1987	1990	1992	1995
AGR	0.0740	0.0655	0.0740	0.0675	0.0740	0.0618	0.2812	0.3323	0.2812	0.3065	0.2716	0.3386
MTG	0.0063	0.0072	0.0063	0.0057	0.0052	0.0055	0.9379	0.9240	0.9379	0.9402	0.9476	0.9404
PWR	0.0680	0.0624	0.0680	0.0717	0.0691	0.0751	0.2396	0.2513	0.2396	0.2172	0.2235	0.1979
COK	0.0049	0.0053	0.0049	0.0048	0.0047	0.0040	0.9485	0.9418	0.9485	0.9460	0.9485	0.9514
OIL	0.0669	0.0670	0.0669	0.0612	0.0682	0.0623	0.2842	0.2290	0.2842	0.2523	0.2341	0.2118
CHM	0.0121	0.0083	0.0121	0.0110	0.0079	0.0075	0.8746	0.9108	0.8746	0.8722	0.9120	0.8993
MEC	0.0667	0.0671	0.0667	0.0568	0.0637	0.0470	0.3152	0.2860	0.3152	0.3219	0.2418	0.2775
CMI	0.0053	0.0048	0.0053	0.0080	0.0097	0.0117	0.9496	0.9555	0.9496	0.9227	0.9051	0.8479
WDF	0.0720	0.0741	0.0720	0.0684	0.0640	0.0568	0.2150	0.1954	0.2150	0.2401	0.2679	0.3131
FDI	0.0092	0.0085	0.0092	0.0095	0.0074	0.0074	0.8961	0.9088	0.8961	0.8962	0.9165	0.9143
WVI	0.0526	0.0573	0.0526	0.0583	0.0590	0.0652	0.3020	0.2654	0.3020	0.3009	0.2386	0.2068
LTR	0.0071	0.0086	0.0071	0.0056	0.0051	0.0063	0.9096	0.8943	0.9096	0.9304	0.9354	0.9229
PCE	0.0605	0.0626	0.0605	0.0599	0.0604	0.0639	0.2417	0.2413	0.2417	0.2279	0.1913	0.2192
OTI	0.0054	0.0047	0.0054	0.0049	0.0049	0.0031	0.9399	0.9451	0.9399	0.9449	0.9434	0.9641
CST	0.0620	0.0613	0.0620	0.0653	0.0620	0.0650	0.2494	0.2308	0.2494	0.2302	0.2314	0.2660
TRN	0.0709	0.0582	0.0709	0.0685	0.0568	0.0552	0.2288	0.3751	0.2288	0.2446	0.3422	0.3976
BFI	0.0560	0.0211	0.0560	0.0185	0.0260	0.0175	0.4430	0.7894	0.4430	0.8163	0.7365	0.8202
NMP	0.0178	0.0618	0.0178	0.0522	0.0475	0.0676	0.8136	0.3536	0.8136	0.4713	0.4896	0.3253