Analysis of Staffing in Two China Government Agencies

Jenne Hui — Katie Li — Harry Xu

STAT 450

Abstract

This project aims to understand the distribution and trends of employment for two of China's civil government branches: Taxation Administration Services (TAS) and Environmental Protection Agencies (EPA). To better visualize the data, plots were generated to show how the composition of staff for different education and work unit levels has changed over time. A statistical test was performed to test if any economic or demographic variables had significant effects on staffing numbers. It was found that public expenditure and GDP per capita are consistently significant predictors of staffing for many variables. A model-based clustering analysis forms homogeneous groups of provinces that have similar trends in the proportion of staff at each variable level. Several groups of provinces were found to be grouped together for most of the levels of age and education levels, while other provinces had more unique trends.

1. Introduction

This report presents graphical, clustering, and regression analyses of employment levels for two of Chinas civil government branches: Taxation Administration Services (TAS) and Environmental Protection Agencies (EPA). Civil services in China are very decentralized and exist on several levels, from provincial to city.

Several years of data on staffing numbers of these two government agencies were used to analyze trends and similarities between provinces and other factors of interest. The contents of this report explore several domains of analysis. Graphical visualization techniques were used to generate meaningful ways to visualize trends and changes in government staffing. Clustering analysis was used to determine which provinces have similar variation in staffing both between and within the two agencies. Additionally, panel regression analysis was performed to determine significant variables that explain the variation in number of staff per province per time.

The report is structured as follows: Section 2 is an overview of the datasets used throughout the case. Section 3 explains the visualization and statistical analyses methods. Section 4 summarizes and interprets the results of the analysis. Lastly, Section 5 provides concluding remarks about the project.

2. Data

Employment data for the two civil service branches (TAS and EPA) were separately hand-collected by research students as part of a study. Tables 1 and 2 show an overview of

Table 1: Summary of TAX dataset

Variable	Levels	Years Spanned
Age	8 levels (<30, 31-35, 36-40, 41-45, 46-50, 51-54, 55-59, >60)	1996, 2000-2007
Education	7 levels (Junior High, High School, Technical College, Other Post-Secondary, Bachelors, Post-Graduate)	1996, 2000-2007
Work Unit	4 levels (Admin, Directly Affiliated (DAE) Admin, Tax, Directly Affiliated (DAE) Non-Admin)	2000-2003
Bureau	2 levels (STB and LTB)	1996, 2000-2007

Table 2: Summary of EPA dataset

Variable	Levels	Years Spanned
Department	7 levels (EPA, Monitoring, Inspection, Research, Education, Information, Other)	1992-1993, 1995-2001, 2004-2011
Government Level	4 levels (Provincial, Prefectural, County, Township)	

the datasets contents. While both datasets are structured as panel data and contain staffing counts for each of Chinas 30 provinces, they differ in the categorizations of staff and the number of years spanned. For the TAS dataset[1], the staff numbers are categorized by age, education, and work unit levels. They are also separated into two bureaus - state tax bureau (STB) and local tax bureau (LTB). On the other hand, the EPA dataset has staffing counts categorized by department and the level of government.

A third dataset with several economic and demographic variables at the provincial level was also provided by the client. The data were extracted from economic journals. These economic and demographic indicators were used in panel regression to determine if some variables partially relate to TAS and EPA staffing. The contents of this dataset are summarized in Table 3.

Table 3: Summary of econ and demographic variables dataset

Variable	Description	Years
FDI	Foreign direct investment	2000-2013
urbanarea	Urban Area	
No.of.above.size.industrial.companies	No. Above-Size Industrial Companies	
public.expenditure	Public Expenditure	
GDP	GDP	
public..employment.thousands	Public Employment (In Thousands)	
Population	Population	
RatioofUrbanPopulation	Ratio of Urban Population	
GDPperCapita	GPD per Capita	
Areasquarekm	Area in KM^2	

[1]The client provided both the original TAS dataset and a partially-cleaned version in which he imputed data from 2005 and 2006 to account for temporary workers. In our analysis, we use the partially-cleaned data.

3. Methodology

Visualization

Tableau, a visualization software, was extensively used to plot the data and results of the clustering analysis. The ability to group, filter, and select variables was utilized throughout the visualization and analysis process. The user is also able to create calculated fields such as ratios, percentage differences, and proportions of total, which is useful for manipulating data. Tableau is also used to visualize the clustering of Chinas provinces[2]; each color represents a group of provinces singled out by the clustering algorithm, which is explained in the next subsection.

Functional Data Clustering

Functional clustering analysis is used to form homogeneous groups of provinces based on their movement patterns in staffing. First, data are converted into matrices, one for each variable of interest. Each cell of the matrix then corresponds to the observed value of the variable of interest (e.g. total number of staff) in one year (row) and in one province (column). Year-by-province matrices were constructed for the following variables: total number of staff, proportion of staff in a given level of age/education/work unit/department, and year-over-year percentage change in staffing.

If we cluster on the total counts of staff, the algorithm will determine cluster membership by overall height or amplitude of the curve, as evidenced in Figure 1. Because there is nothing informative about these clusters, we convert staffing counts into percentages to reduce the amplitudes and extremes in the functions that represent the staffing changes over the years. This helps focus the clustering algorithm away from inherent differences in staffing magnitude between each province.

Each provinces observed staffing counts over the years is treated as a function. Given the data matrix, we then convert the data into functional form using curve-smoothing techniques. Lastly, a model-based clustering algorithm is applied to group provinces based on a predetermined number of clusters. The details of the statistical software package used can be found in Appendix section 1.

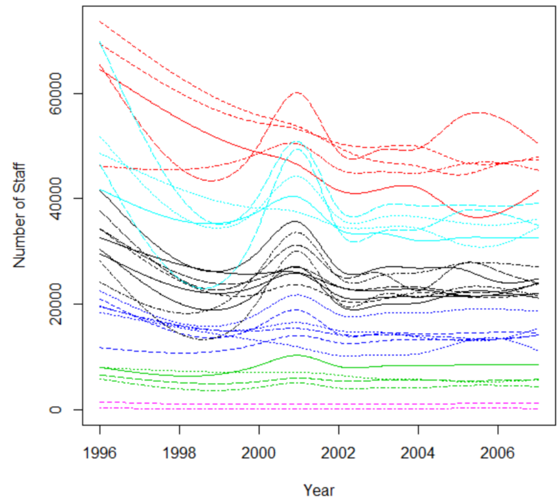
Panel Regression

To determine which variables are significantly associated with staffing counts, a two-way fixed-effect panel regression model is applied to the data. A panel regression model is necessary because staffing is observed for each province at multiple time points (years), rather than one point in time. A panel linear model differs from traditional linear models because it controls for unique characteristics across provinces and years so that the effects of the explanatory variables on staffing can be extracted.

Individual effects are unobserved or unmeasurable variables that are fixed in time but vary across provinces. Time effects are variables that occur across all provinces but differ from year to year. The model was created using two-way fixed effects to control for both of these effects. We are not interested in time-invariant variables and variations within provinces are

[2] Latitudinal and longitudinal data were sourced from https://tableaumapping.wordpress.com/2013/12/17/china-provinces/

Figure 1: Total number of staff for each province color-coded by cluster membership (TAS dataset)



assumed to be large and possibly correlated with regressors; hence, a fixed individual effect is necessary. The Hausman test is applied to check this fixed model assumption. The results of the test show that the fixed-effect model is preferred to the random-effect model. The test for individual and time effects shows significance in both effects. Therefore, using a fixed two-way effect model should be sufficient.

From the 10 economic variables given, 5 variables are dropped due to missing values and possible multicollinearity. For example, GDP per capita is collinear with GDP and population as the former is just a transform of the latter two variables. The 5 variables that are incorporated into the model are urban area, number of above-size industrial companies, public expenditure, GDP per capita, and number of public employees. The variable number of public employees is shown to not be a significant independent variable for all sub-level models, so it is disregarded.

The mathematical details of the fixed-effect model and diagnostic tests are described in Appendix section 2.

4. Results

Visualization Results

To look at the composition of staff based on age, education and work unit, we plotted the staff count for each level (as a percentage of total) as a stacked line graph to see how the distribution of age, education, and unit varied over time for each bureau.

Figure 2 shows the composition of staff based on education level; the staff counts are summed across all provinces. The percentage of total staff that had a Bachelors degree grew rapidly, from only 5.6% in 1996 to 46% in 2007. On the other hand, the percentage of staff with a Technical college background consistently decreased each year. The education group with the largest percentage of staff is Other Post-Secondary education - it grew from 32% in 1996 to a maximum of 53% in 2003 but then decreased yearly, reaching 40% in 2007.

Figure 3 shows the composition of staff based on work unit. There is one plot for each bureau, LTB and STB. In 2000, the Tax Office unit had the most staff, with 63%, however, this percentage decreased yearly and in 2003, the percentage of staff working in a tax office was only 34%. This decrease was most pronounced in the STB bureau group, as the percentage of STB staff in Tax offices decreased from 62% to 27% - however this decrease in STB was made up for by an increase in DAE Admin and DAE Non-Admin workers.

It was also of interest to see if these trends were the same for every province. To gain insight into provincial variation we replicated these charts for all provinces. For age groups, it is immediately clear that Tibet's workers have a much higher proportion of workers under the age of 30 years, compared to other provinces. When it comes to Education group, Hainan and Yangzhou greatly vary from the other provinces. Hainan has a much higher proportion of high school workers for most years. While most provinces predominantly consist of "other post-secondary" workers, there are very few of them in Yangzhou - instead there is a much higher proportion of post-graduate workers.

Figure 2: Distribution of TAS staff by Education Level

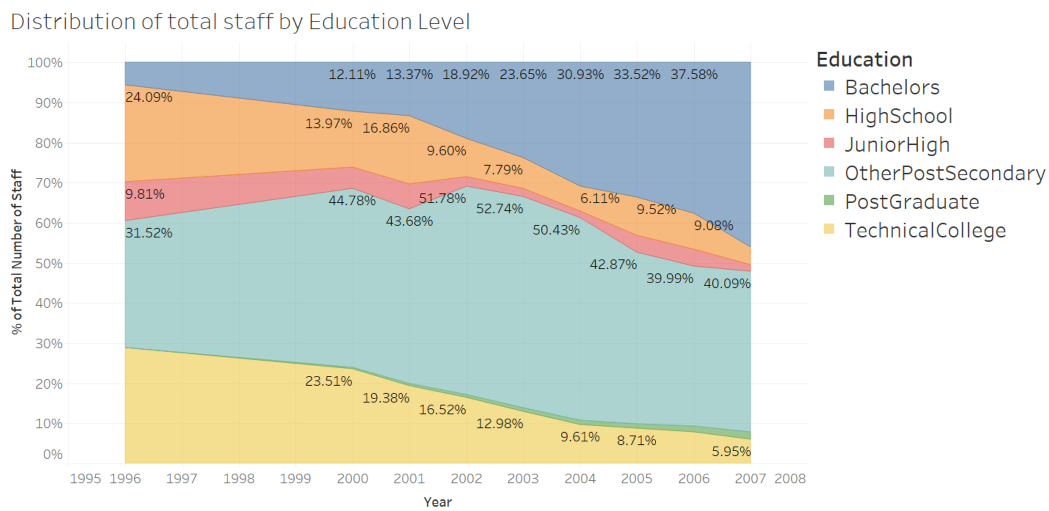
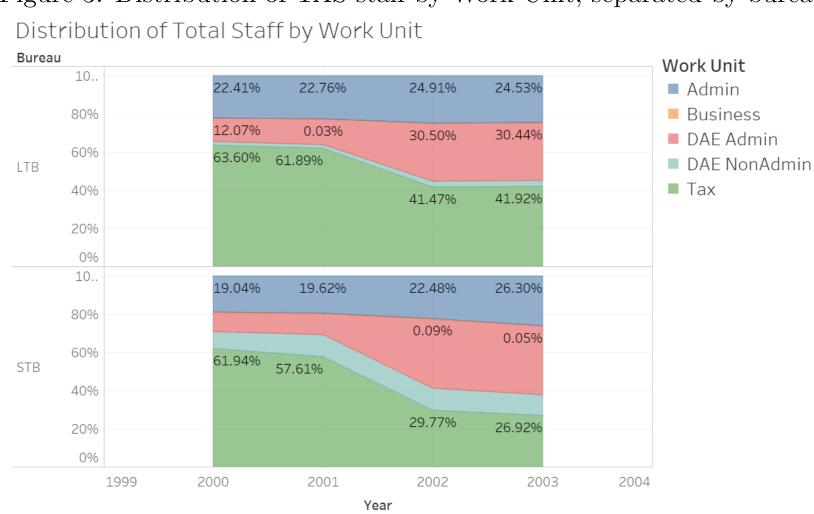


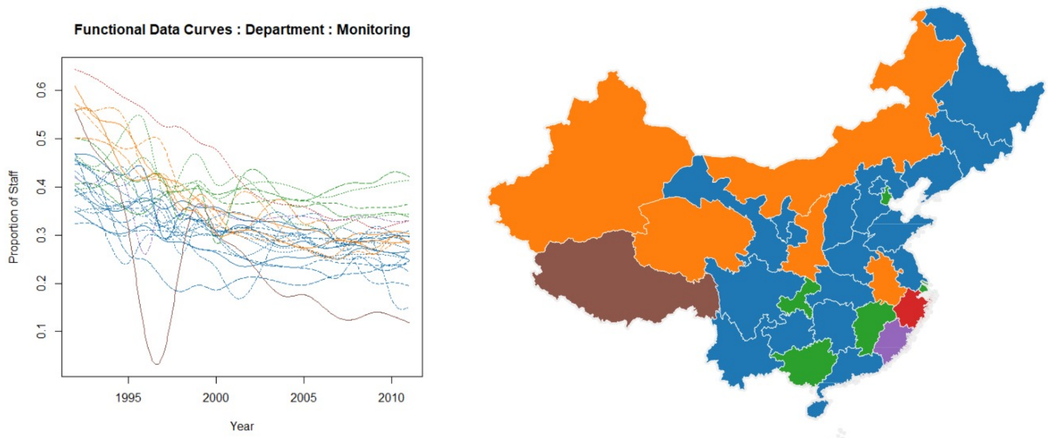
Figure 3: Distribution of TAS staff by Work Unit, separated by bureau



Clustering Results

As mentioned in the previous section, the objective of clustering is to determine groups of provinces that have similar movement patterns in staffing. As an example, Figure 4 shows the results for 6 clusters on the proportion of staff in the Monitoring department in the EPA agency. On the left are smoothed provincial curves and on the right is the map of China. Each color represents a different cluster (group of provinces).

Figure 4: Functional data curves and corresponding geographic cluster plot for the Monitoring department in the EPA agency



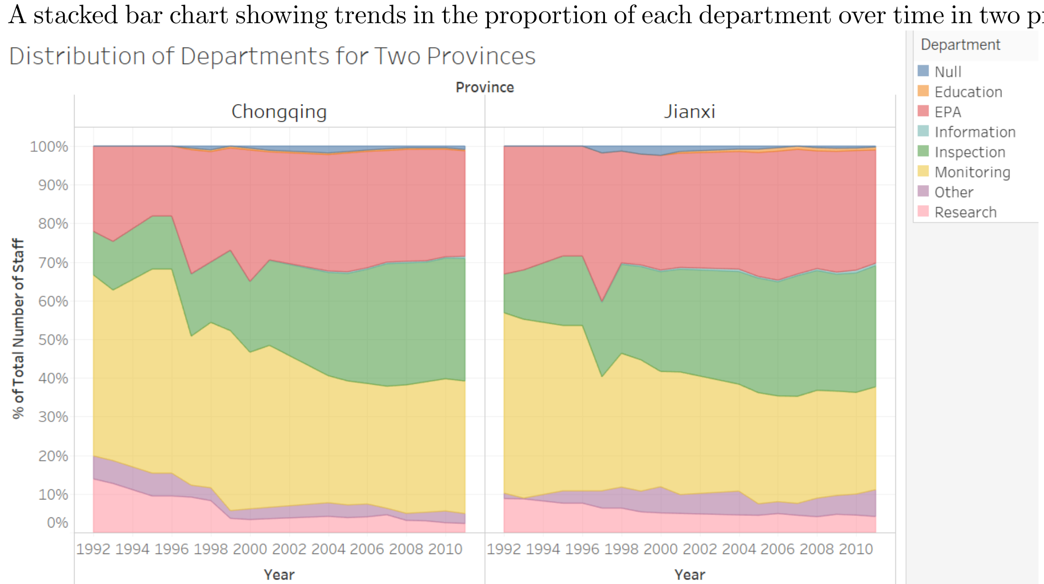
In the TAS dataset, staffing was broken down into different age groups and education levels. Due to significant amount of missing data from 1996 to 1999, we reduced our analysis to data between 2000 to 2007. All of the cluster plots can be found in Appendix sections 4 and 5. The two main objectives of the analysis were to determine how provinces are clustered for each of the different age and education levels and if there were any sets of provinces that were recurrently grouped together across most levels. A similarity matrix for each of age and education can be found in Appendix section 6. Each cell in the matrix corresponds to a pair of provinces, and the value represents the number of levels in which the provinces were clustered together. For example, since there are 8 age groups, the largest possible entry in the age similarity matrix is 8.

At first glance, we notice that the matrix contains relatively few 0 entries and its predominant entry is 1. This implies that almost all pairs of provinces are grouped together at least once across the different age levels. The column sums are presented in the bottom row. The largest of these values belongs to Guizhou (100), Guangxi (99), Gansu (99), Hunan (99), and Yunnan (99), meaning those provinces often occur in large, recurrent clusters. On the other hand, Yangzhou has a strikingly low value of 28. This low values implies that Yangzhou does not share many clusters with the other provinces, as its trends across the age groups are unique in comparison with other provinces. In addition, Guizhou and Yangzhou are the only provinces that were paired together in clusterings for all eight age levels. There are also a moderate number of other province pairs that are paired together in clusters for seven out of eight age levels, such as Sichuan and Guizhou, and Hubei and Shandong.

In contrast to the results from the age group analysis, the results from the education analysis implies that provinces are not very similar across the six levels. The predominant values in the education similarity matrix are 0 and 1. This means that certain provinces may share the same cluster for one education level, but not for the others. Moreover, we do not see any distinct provinces that are paired together in clusterings across the six levels. There are only two province pairs that shared 5 clusters, namely Shandong and Guangdong, and Inner Mongolia and Yunnan.

For the EPA dataset, staffing numbers were broken down by department and level of government rather than by age and education. Two provinces, Hunan and Yunnan, were not included in the analysis because data was missing for some years. After clustering the provinces for each of the seven departments and four government levels, we again performed a similarity analysis to see if some provinces appeared in the same clusters for most of the departments. We found that Chongqing and Jiangxi province were clustered together in 6 out of 7 departments, as well as Guangdong and Shandong. Figure 5 shows the distribution of departments for Jiangxi and Chongqing - the trends in the composition of staff are very similar. For the full cluster results, refer to Section 4 of the Appendix.

Figure 5: A stacked bar chart showing trends in the proportion of each department over time in two provinces.



For staffing at the four levels of government, it was found that Liaoning and Gansu were clustered together for all four levels. Shandong and Shanxi also appeared in the same cluster for all levels, meaning that the trends of staffing for those provinces are very similar. Provinces that were clustered together in three out of four levels are summarized in Table 4. Due to the large amount of missing data for the Township level, only a few provinces were able to be included in the cluster analysis for that level. Tibet was the only province to appear in its own cluster for two out of four levels, Provincial and Prefectural, which shows that it exhibited unique trends in staffing.

Panel Regression Results

Panel regression was applied to total staff counts for the TAS dataset. When combining LTB and STB bureaus, two out of four model variables were determined to be significant, GDP per capita and public expenditure. However, when fitting a model for the bureaus separately, all four variables were significant for STB while none were significant for LTB.

Table 4: Groups of provinces that were clustered together in three out of four government levels for the EPA dataset

Government Levels Clustered Together	Provinces in Cluster		
Provincial Prefectural County	Anhui Gansu Guangdong Guangxi Inner Mongolia Liaoning	Chongqing Fujian Heilongjiang Jiangsu Jiangxi Zhejiang	Henan Jilin Shandong Shanxi
Provincial County Township	Hebei Shandong Shanxi		
Prefectural County Township	Xinjiang Gansu Liaoning		

Table 5: Coefficients of significant regressors in panel regression model on total staff

Combined STB and LTB	STB only	LTB only
GDP per capita: 0.17	GDP per capita: 0.076	None
Public expenditure: -3.37	Public expenditure: -2.40	
	Urban area: 1.56	
	# above-sized industrial companies: -0.077	

The coefficients for these models are shown in Table 5. For the combined STB and LTB data, we can interpret the contribution of public expenditure as a one unit increase in public expenditure is associated with a decrease in 3.37 units in total staff count. In addition, each province and year has its own effects that contributes to the total staff count in a particular province and year.

Additionally, a panel linear model was fitted for each level of age, education, and work unit for each of the two TAS bureaus. The table in the appendix 2 summarizes the results of the fixed-effects regression, with each models significant regressors. Generally, it seems that GDP per capita and public expenditure have a consistently strong association with staffing counts. Especially for the education variable, GDP per capita showed up as a significant regressor in 9 out of 12 models (six levels of education in two bureaus). Another interesting finding was that for the age groups 55-59 and 60 and above, none of the regressors were significant predictors of staff counts. This was the case for both tax bureaus.

5. Conclusions and Recommendations

As evidenced by the analyses, a lot of provinces have similar trends in staffing and government composition. However, there is still a lot of provincial variation. In further analysis, it would be useful to explore what drives the similarity within groups as well as the variation between groups of provinces. The interpretation of the clustering results and visualization can be greatly improved if supplemented by additional knowledge of social and policy changes in China for the applicable years of data.

For now, we can try to reconcile our clustering results with how the provinces are separated into Chinas administrative divisions. This may be a natural way of looking at provincial similarities and differences. As of 2018, China has four main divisions at the provincial level: 1. Provinces, 2. Autonomous Regions, 3. Municipalities, and 4. Special Administrative Regions. Municipalities tend not to be clustered often with other provinces. This is not the case with Chongqing municipality, as the similarity sums (column sums in the similarity matrices) are frequently at the average or high-average range, indicating high similarity with other provinces. A future analysis can focus on investigating how staffing for Chongqing is different from that for the other municipalities, or if there are properties of staffing changes that are characteristic of the municipality provinces. Regarding provinces in the autonomous region (AR) division, we dont see exceptionally high or low similarity sum values with the exception of Tibet. See Appendix section 6 for tables of similarity sums for all provinces as well as the breakdown of provinces into administrative divisions.

Improvements in the both clustering analysis and panel regression analysis approaches can be pursued in the future work on this data. There are some limitations with the panel regression model in the context of our data. Many of the adjusted R-squared levels are quite low, which suggests that the explanatory variables chosen in the fitted model are not explaining much of the variation in staffing counts. Another concern with the panel models is that they do not consider the relationship between the levels in age, education, and unit themselves. For example, the staffing of the 55-60 age group may be strongly correlated with the 60 and above group. Our model could not take dependencies between levels into account; the best we can do is run a regression for each level independently.

6. Appendix

6.1. Functional Clustering in R

To do the clustering, we used the "FunFEM" package in R. Before this, we tried the 'Funclustering" package but found that the results were more stable and satisfactory with "FunFEM". The reader should be advised that we found a bug in the "FunFEM" source code. If you would like to reproduce the results of this analysis and are having trouble with the "FunFEM" package, feel free to contact `katieli16@yahoo.com` for the code fix.

6.2. Two-way fixed-effect panel regression

The two-way fixed-effect panel model used can be expressed mathematically as:

$$Y_{it} = \beta_1 X_{it} + \alpha_i + \lambda_t + \epsilon_{it}$$

where Y_{it} is the response variable representing the number of staff for a specific province i in a specific year t. The matrix X_{it} are the possible economic explanatory variables that could explain the response variables, and β_1 are the slopes for the explanatory variables of interest for each X_{it}. α_i is the entity (Province) fixed effect, and λ_t is the time fixed effect. Both terms have a specific effect on Y_{it}. In this model, time t takes value 1996, 2000 to 2013, and ϵ_{it} is the error term for Province i in year t. It is assumed that the errors are iid and $\epsilon_{it} \sim N(0, \sigma^2)$.

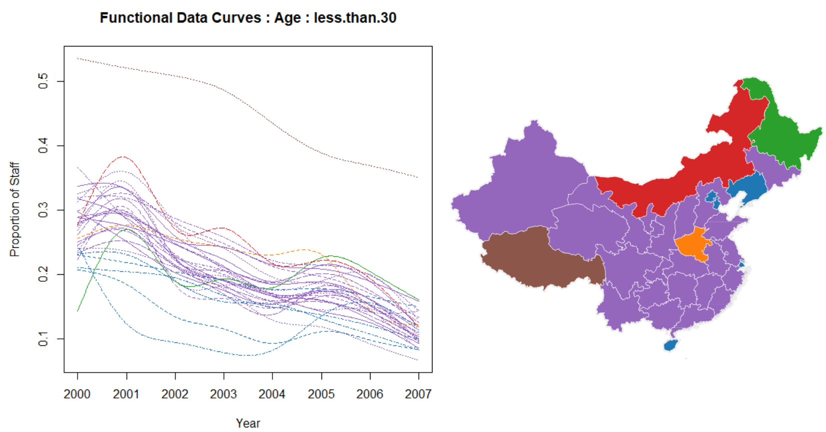
6.3. Significant Regressors in Panel Linear Models on TAS dataset variables

Response Variable (Staffing Counts)	**STB Significant Regressors**	**LTB Significant Regressors**
Age Under 30	GDP per Capita GDP Public Expenditure	GDP per Capita GDP Urban Area
Age 30-35	GDP per Capita GDP Urban Area Public Expenditure	GDP Public Expenditure
Age 36-40	GDP per Capita GDP Number Above Size Industrial	GDP per Capita Urban Area Public Expenditure
Age 41-45	GDP per Capita Public Expenditure	GDP per Capita Public Expenditure Urban Area
Age 46-50	None	Urban Area
Age 51-54	GDP per Capita GDP Urban Area Number Above Size Industrial	GDP per Capita GDP Urban Area Number Above Size Industrial Public Expenditure
Age 55-59	None	None
Age 60 and Above	None	None
Edu Post-Graduate	GDP per Capita Urban Area Public Expenditure	Urban Area Public Expenditure
Edu Bachelor's	GDP per Capita GDP Public Expenditure	GDP per Capita Number Above Size Industrial
Edu Other Post-Secondary	GDP Urban Area Number Above Size Industrial	GDP
Edu Technical College	GDP per Capita Public Expenditure	GDP per Capita Public Expenditure
Edu High School	GDP per Capita	GDP per Capita Urban Area
Edu Junior High	GDP per Capita	GDP per Capita Above Size Industrial
Unit Admin	None	None
Unit Directly Affiliated Admin	GDP per Capita GDP	GDP per Capita Number Above Size Industrial
Unit Directly Affiliated Non-Admin	GDP Number Above Size Industrial	None
Unit Taxation	GDP per Capita GDP	GDP per Capita Public Expenditure

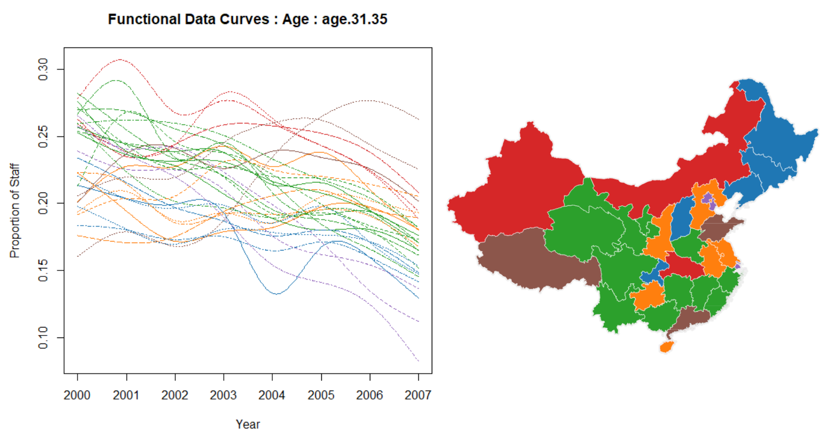
6.4. Cluster Results for TAX Dataset

For each level of Age and Education, the results of the clustering as well as the cluster memberships are shown. The response variable that is being clustered is the proportion of staff at the given variable level.

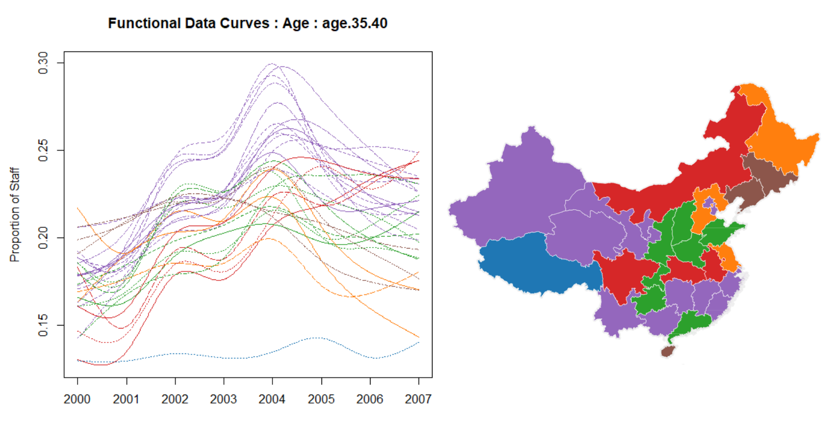
Age 30 and below



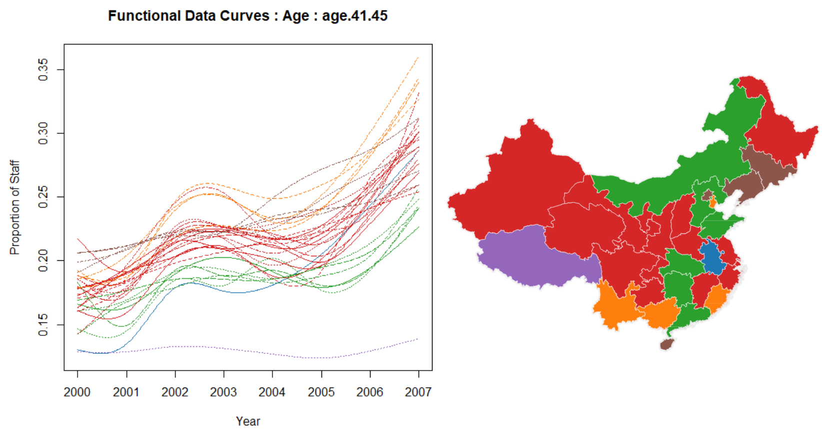
Age 31-35



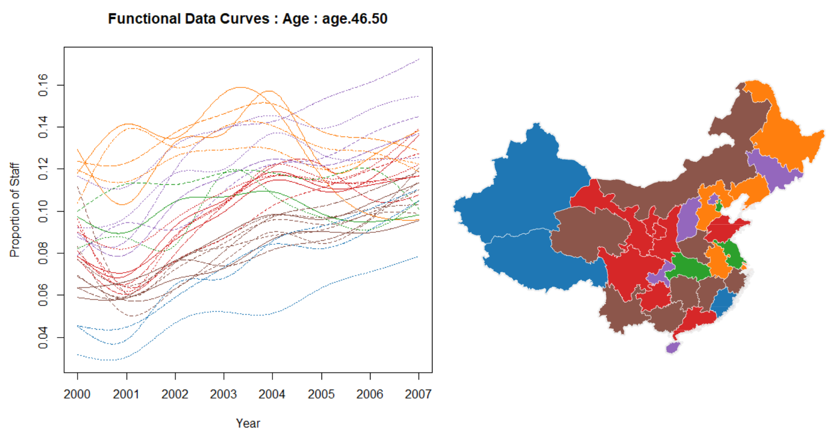
Age 35-40



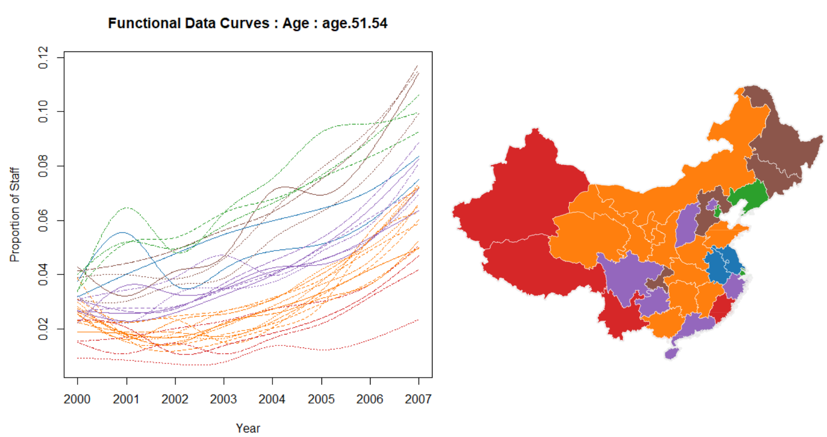
Age 41-45



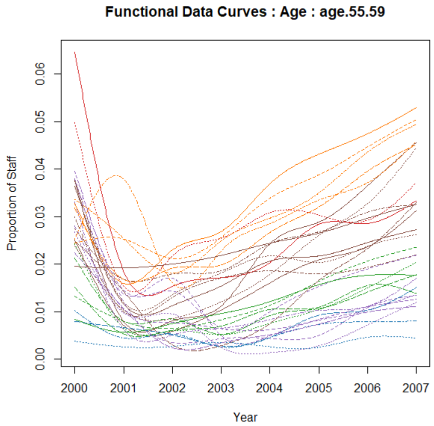
Age 46-50



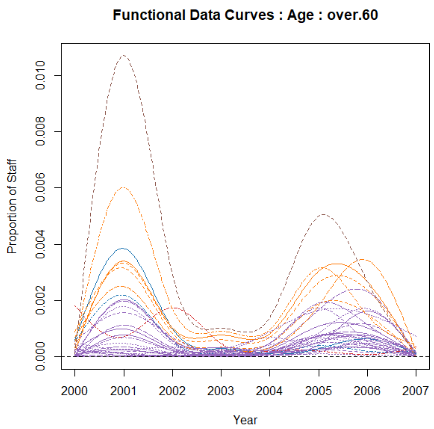
Age 51-54

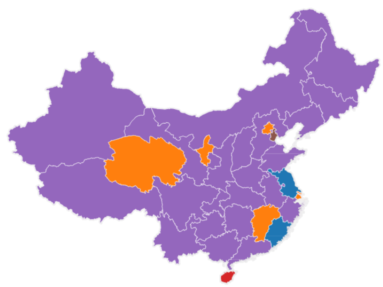


Age 55-59

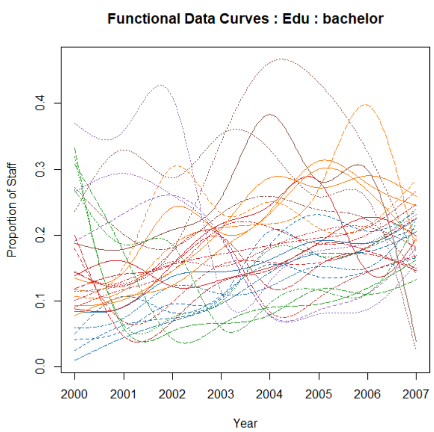


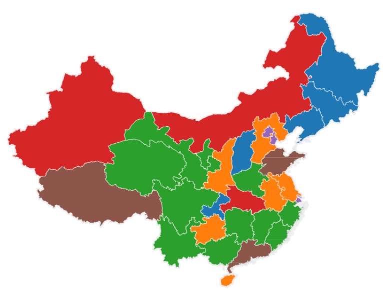
Age 60 and above



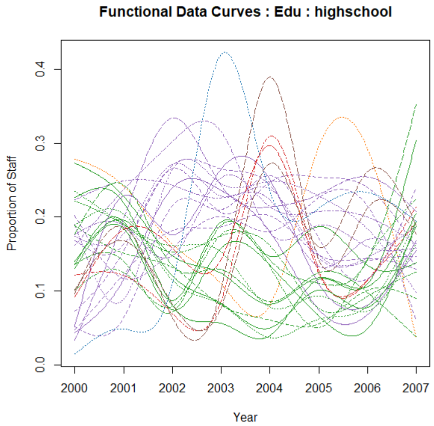


Bachelors Education

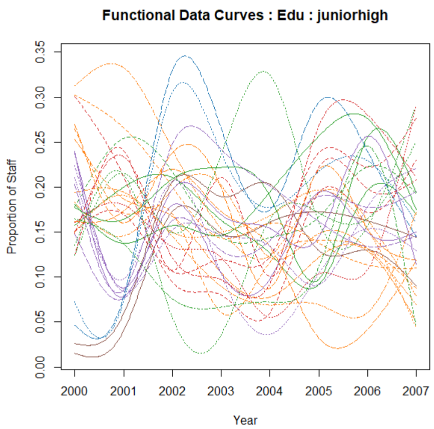




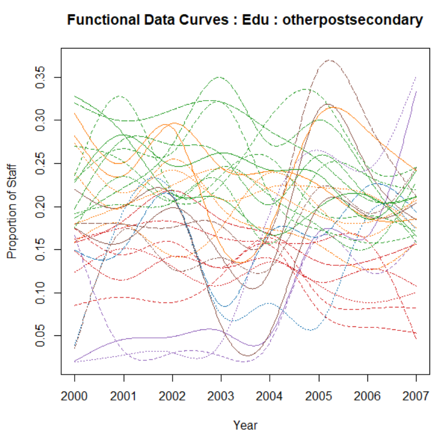
High School Education



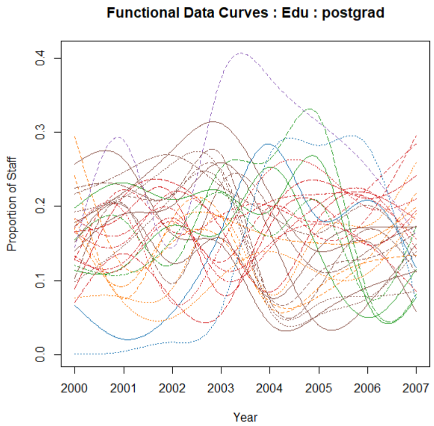
Junior High Education

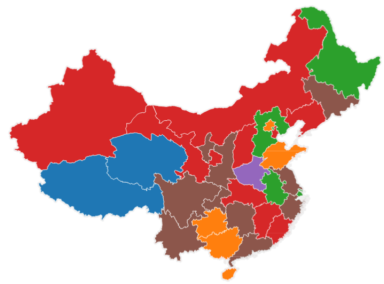


Other Post-Secondary Education

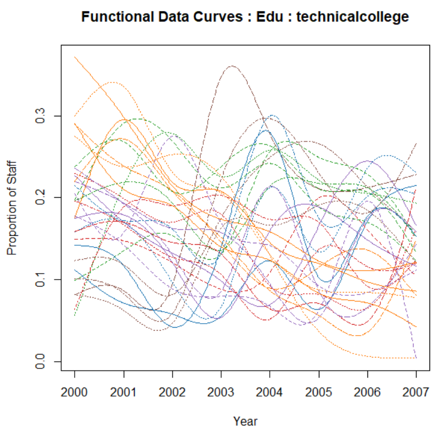


Post-Graduate Education





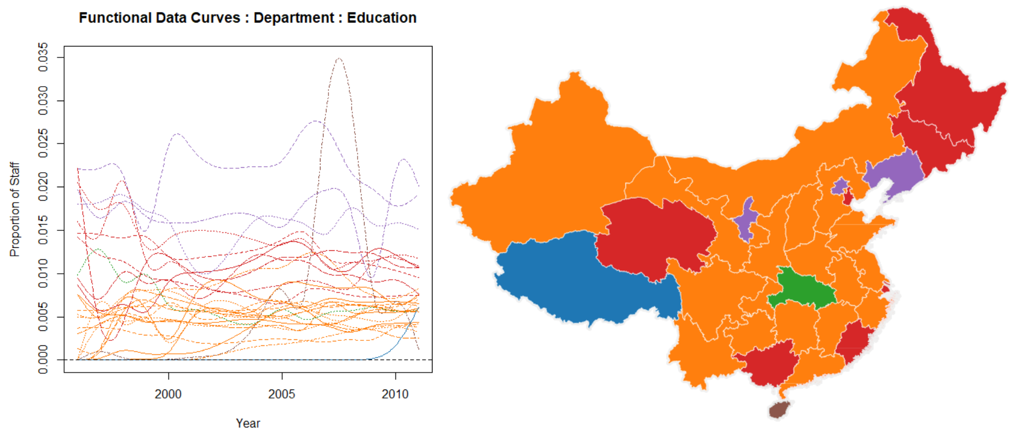
Technical College Education



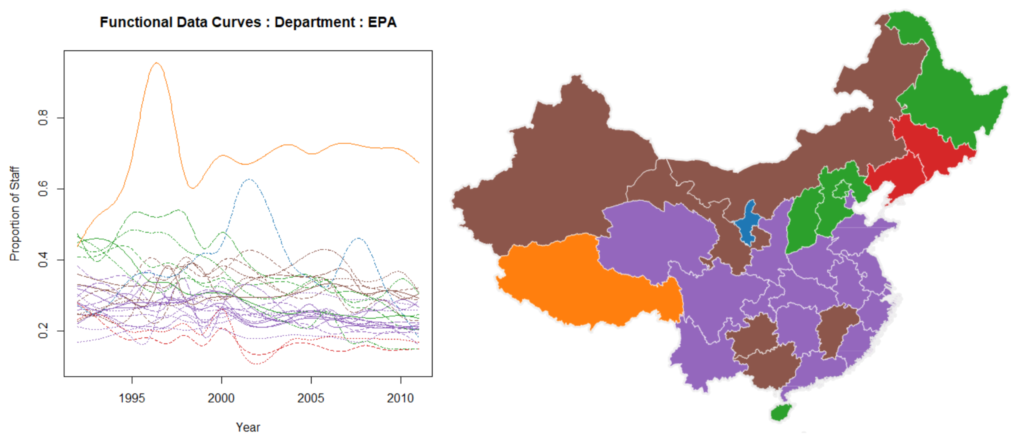
6.5. *Cluster Results for EPA Dataset*

For each level of Department and Gov't Level, the results of the clustering as well as the cluster memberships are shown. The response variable that is being clustered is the proportion of staff at the given variable level.

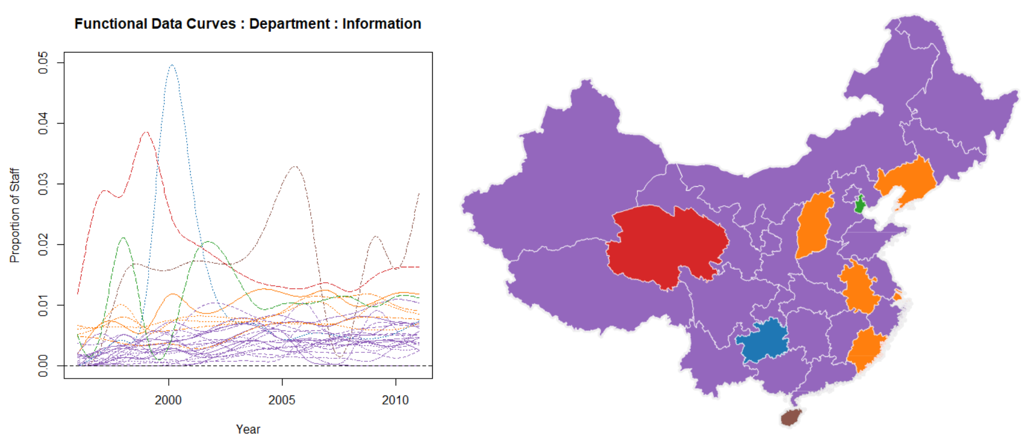
Education Department



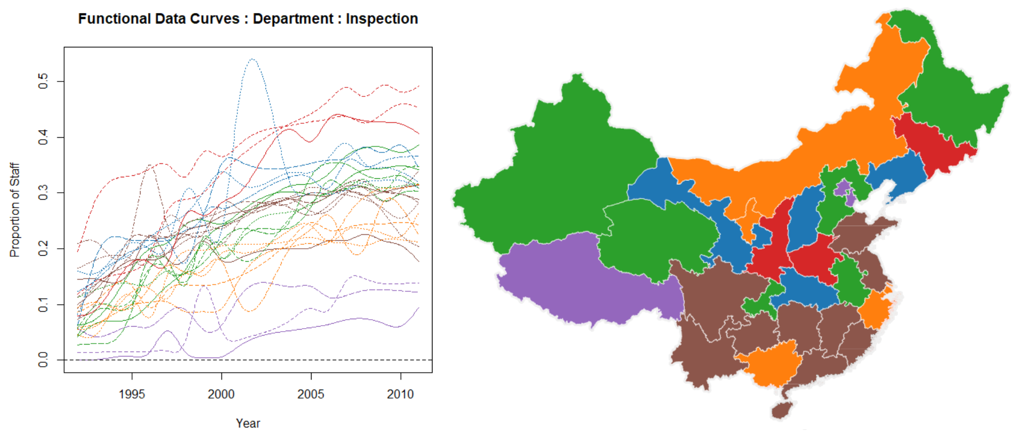
EPA Department



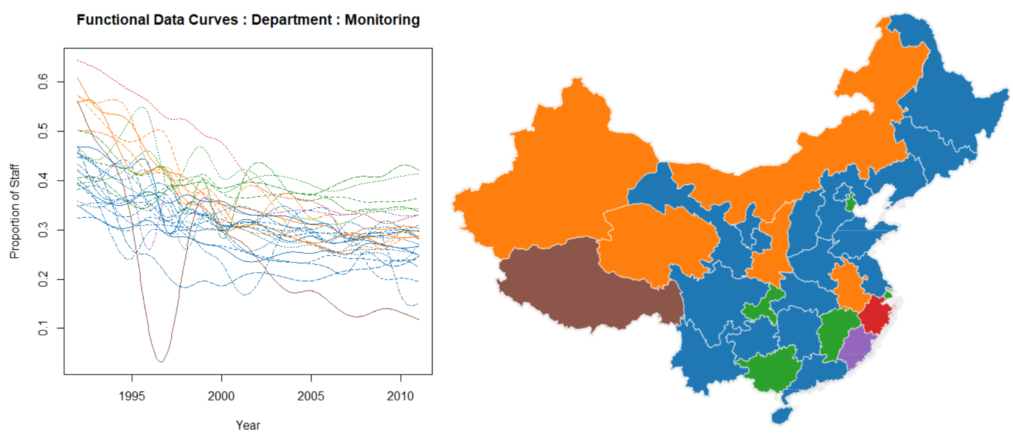
Information Department



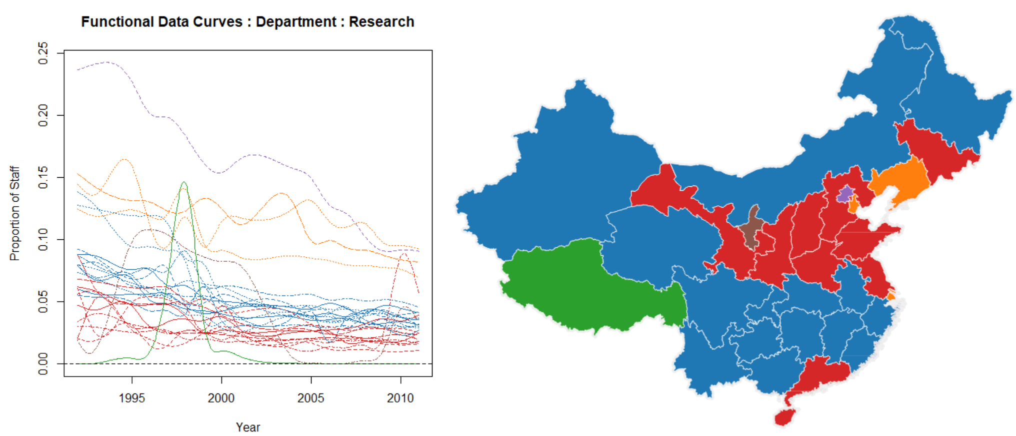
Inspection Department



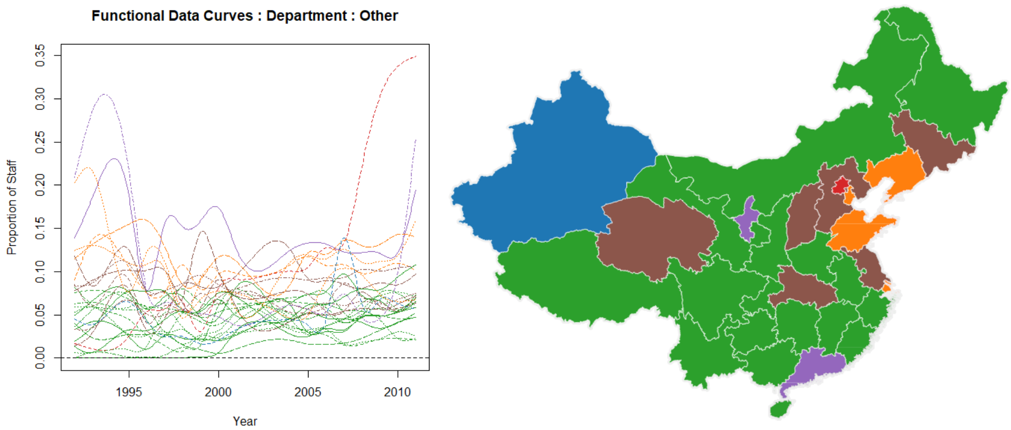
Monitoring Department



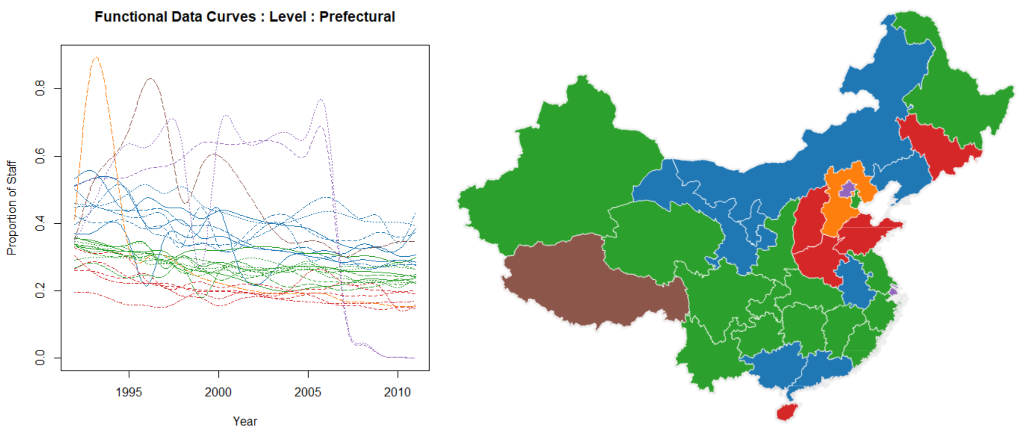
Research Department



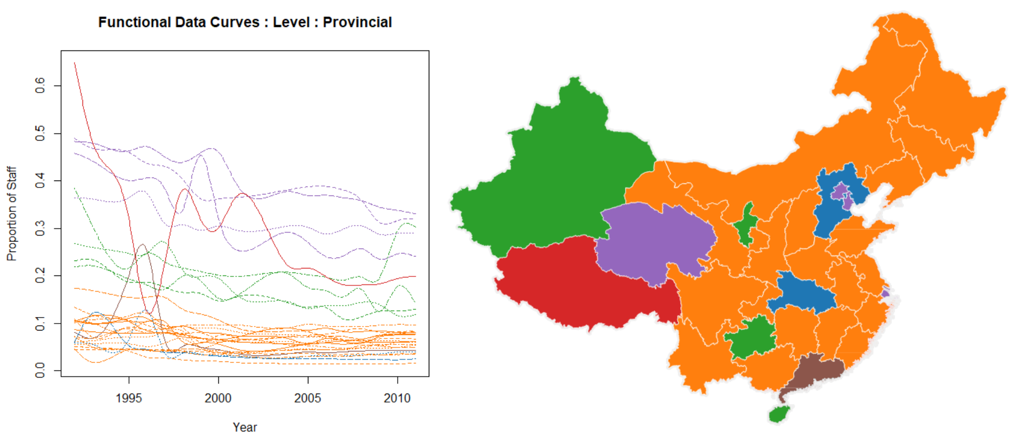
Other Department



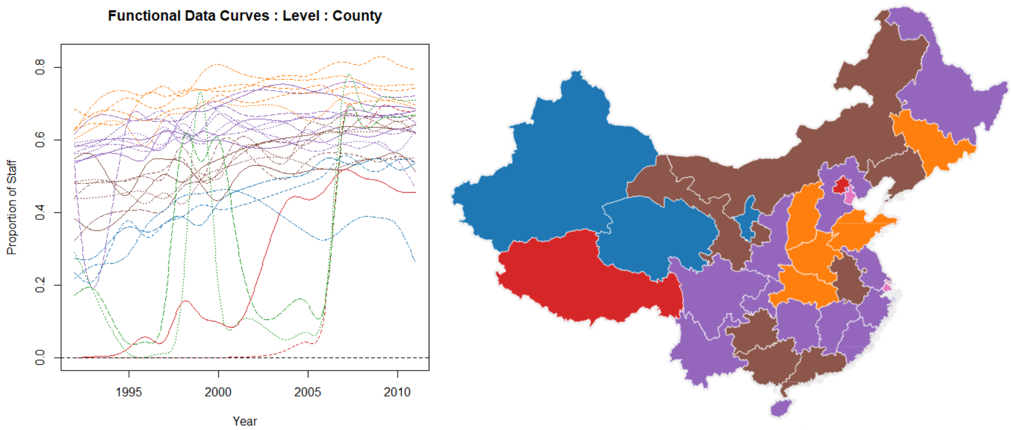
Prefectural Level



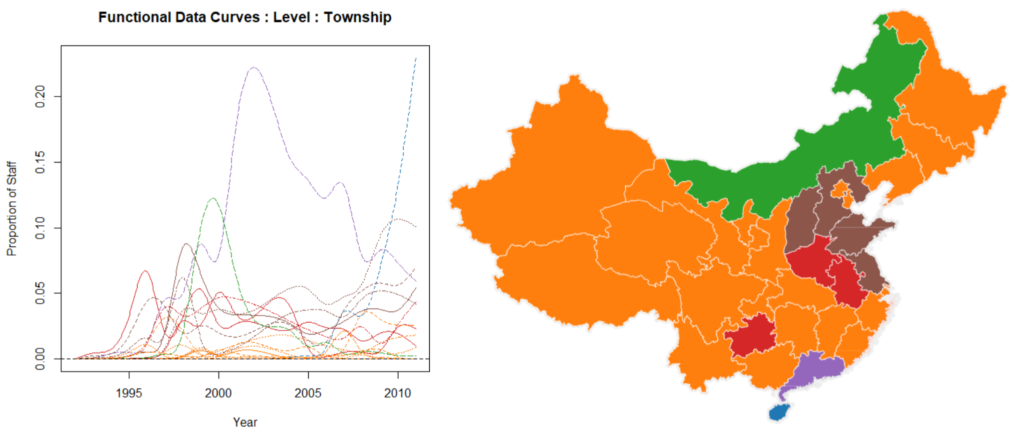
Provincial Level



County Level



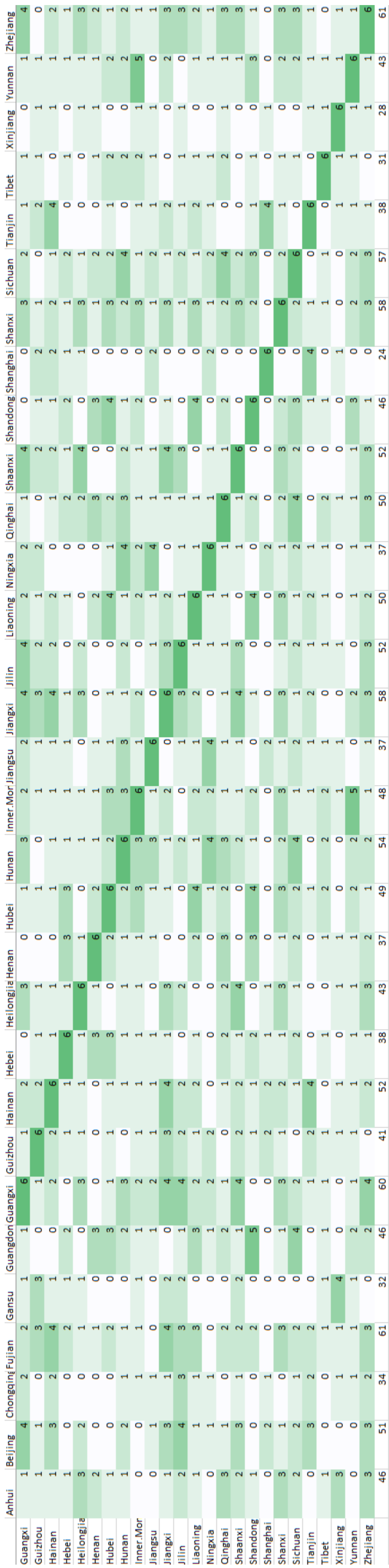
Township Level



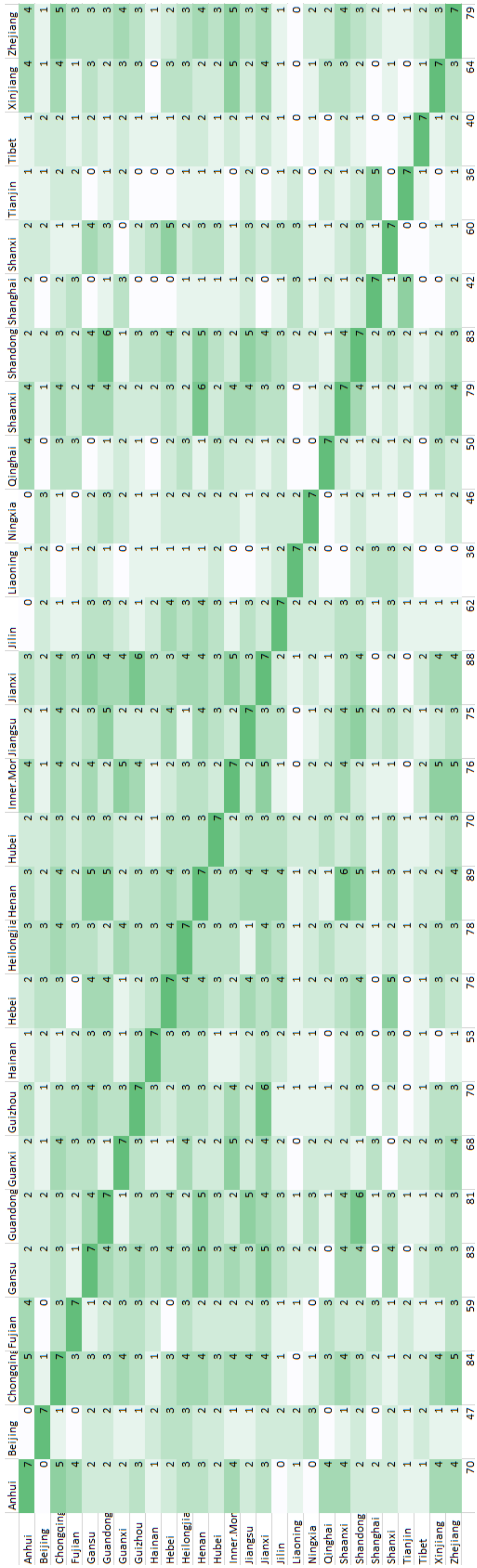
6.6 - Similarity Matrices

a. TAX : Age

	Anhui	Beijing	Chongqir	Fujian	Gansu	Guangdo	Guangxi	Guizhou	Hainan	Hebei	Heilongji	Henan	Hubei	Hunan	Inner.Mo	Jiangsu	Jiangxi	Jilin	Liaoning	Ningxia	Qinghai	Shaanxi	Shandon	Shangha	Shanxi	Sichuan	Tianjin	Tibet	Xinjiang	Yunnan	Zhejiang
Anhui	9	0	4	2	3	3	3	4	1	5	2	2	4	3	3	4	2	2	2	2	2	4	3	1	3	4	0	1	2	3	3
Beijing	0	9	1	1	2	1	1	1	5	0	0	1	1	0	0	0	2	3	3	3	3	0	0	4	2	1	4	0	1	1	2
Chongqir	4	1	9	2	4	4	3	5	1	4	4	3	3	4	2	3	3	5	2	3	3	5	4	1	7	4	0	1	3	3	4
Fujian	2	1	2	9	4	3	5	3	1	2	1	3	2	3	1	3	4	2	0	4	4	2	2	1	3	4	2	2	4	6	5
Gansu	3	2	4	4	9	4	6	5	0	3	2	6	5	5	3	3	6	2	1	8	7	6	5	2	4	6	1	1	4	5	6
Guangdo	3	1	4	3	4	9	3	7	2	4	2	2	4	5	3	2	2	3	1	3	2	5	7	0	6	6	0	2	2	3	5
Guangxi	3	1	3	5	6	3	9	3	0	3	1	7	4	7	4	2	7	2	1	5	6	5	5	1	3	4	2	1	3	8	6
Guizhou	4	1	5	3	5	7	3	9	3	4	3	3	3	4	2	4	3	3	1	4	3	7	5	1	7	7	0	1	3	3	6
Hainan	1	5	1	1	0	2	0	3	9	1	1	0	0	0	0	1	0	5	4	0	0	1	0	1	3	2	2	0	0	0	2
Hebei	5	0	4	2	3	4	3	4	1	9	4	2	4	4	3	5	2	3	3	2	2	4	4	2	3	3	1	1	2	3	3
Heilongji	2	0	4	1	2	2	1	3	1	4	9	2	1	1	1	2	1	4	3	1	1	2	1	2	4	3	0	1	2	1	3
Henan	2	1	3	3	6	2	7	3	0	2	2	9	3	6	4	2	7	1	1	5	6	5	4	2	3	4	1	1	3	6	6
Hubei	4	1	3	2	5	4	4	3	0	4	1	3	9	5	6	3	3	2	1	4	4	4	5	0	3	4	1	1	3	3	3
Hunan	3	0	4	3	5	5	7	4	0	4	1	6	5	9	5	2	6	2	1	4	5	6	7	0	4	4	0	1	2	6	5
Inner.Mo	3	0	2	1	3	3	4	2	0	3	1	4	6	5	9	1	3	1	1	2	3	3	4	0	2	3	0	2	3	3	3
Jiangsu	4	0	3	3	3	2	2	4	1	5	2	2	3	2	1	9	3	1	1	3	3	4	2	2	3	3	2	0	2	2	3
Jiangxi	2	2	3	4	6	2	7	3	0	2	1	7	3	6	3	3	9	1	0	7	8	5	4	3	3	4	1	0	3	6	6
Jilin	2	3	5	2	2	3	2	3	5	3	4	1	2	2	1	1	1	9	5	1	1	2	2	0	5	3	1	1	2	2	3
Liaoning	2	3	2	0	1	1	1	1	4	3	3	1	1	1	1	1	0	5	9	0	0	1	1	4	2	1	4	1	1	1	1
Ningxia	2	3	3	4	8	3	5	4	0	2	1	5	4	4	2	3	7	1	0	9	8	5	4	3	3	5	1	0	3	4	5
Qinghai	2	3	3	4	7	2	6	3	0	2	1	6	4	5	3	3	8	1	0	8	9	4	3	3	3	4	1	0	3	5	6
Shaanxi	4	0	5	2	6	5	5	7	1	4	2	5	4	6	3	4	5	2	1	5	4	9	7	1	5	5	0	1	3	4	4
Shandon	3	0	4	2	5	7	5	5	0	4	1	4	5	7	4	2	4	2	1	4	3	7	9	0	4	4	0	2	2	4	3
Shangha	1	4	1	1	2	0	1	1	1	2	2	2	0	0	0	2	3	0	4	3	3	1	0	9	1	1	5	0	2	1	2
Shanxi	3	2	7	3	4	6	3	7	3	3	4	3	3	4	2	3	3	5	2	3	3	5	4	1	9	6	0	1	3	3	6
Sichuan	4	1	4	4	6	6	4	7	2	3	3	4	4	4	3	3	4	3	1	5	4	5	4	1	6	9	0	1	3	4	7
Tianjin	0	4	0	2	1	0	2	0	2	1	0	1	1	0	0	2	1	1	4	1	1	0	0	5	0	0	9	0	1	2	1
Tibet	1	0	1	2	1	2	1	1	0	1	1	1	1	1	2	0	0	1	1	0	0	1	2	0	1	1	0	9	4	2	1
Xinjiang	2	1	3	4	4	2	3	3	0	2	2	3	3	2	3	2	3	2	1	3	3	3	2	2	3	3	1	4	9	4	4
Yunnan	3	1	3	6	5	3	8	3	0	3	1	6	3	6	3	2	6	2	1	4	5	4	4	1	3	4	2	2	4	9	6
Zhejiang	3	2	4	5	6	5	6	6	2	3	3	6	3	5	3	3	6	3	1	5	6	4	3	2	6	7	1	1	4	6	9
	86	52	104	90	128	105	120	117	45	95	65	110	98	116	80	80	114	79	57	111	112	119	107	55	114	119	42	39	86	113	129



b) TAX : Education



c) EPA : Department

	Anhui	Beijing	Chongqing	Fujian	Gansu	Guandong	Guanxi	Guizhou	Hainan	Hebei	Heilongjia	Henan	Hubei	Inner.Mor	Jiangsu	Jianxi	Jilin	Liaoning	Ningxia	Qinghai	Shaanxi	Shandong	Shanghai	Shanxi	Tianjin	Tibet	Xinjiang	Zhejiang
Anhui	4	0	1	1	3	2	3	2	0	0	1	2	0	3	1	1	1	3	1	0	1	1	0	1	0	0	1	1
Beijing	0	4	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	2	1	0	3	0	2	2	1	0
Chongqing	1	1	4	3	2	0	1	1	1	1	3	1	1	1	3	3	2	2	1	2	4	1	1	1	1	1	1	3
Fujian	1	0	3	3	1	0	1	1	1	1	3	1	1	1	3	3	1	1	0	1	3	1	0	1	0	0	0	3
Gansu	3	1	2	1	4	2	3	1	0	0	1	1	0	3	1	1	2	4	2	1	2	1	1	1	1	1	2	1
Guandong	2	0	0	0	2	4	2	1	0	0	0	0	0	2	0	0	0	2	1	0	0	0	0	0	0	0	1	0
Guanxi	3	0	1	1	3	2	3	1	0	0	1	1	0	3	1	1	1	3	1	0	1	1	0	1	0	0	1	1
Guizhou	2	0	1	1	1	1	1	4	1	0	1	1	1	1	1	1	0	1	1	1	1	0	0	0	0	0	1	1
Hainan	0	0	1	1	0	0	0	1	4	1	1	1	0	0	1	1	1	0	1	0	1	1	0	1	0	0	1	1
Hebei	0	0	1	1	0	0	0	0	1	4	1	0	1	0	2	1	0	0	0	0	1	1	0	1	0	0	0	1
Heilongjia	1	0	3	3	1	0	1	1	1	1	3	1	1	1	3	3	1	1	0	1	3	1	0	1	0	0	0	3
Henan	2	0	1	1	1	0	1	1	1	0	1	4	1	1	1	1	3	1	0	0	1	3	0	3	0	0	0	1
Hubei	0	0	1	1	0	0	0	1	0	1	1	1	3	0	1	1	1	0	0	1	1	1	0	1	0	0	0	1
Inner.Mor	3	0	1	1	3	2	3	1	0	0	1	1	0	4	1	1	1	3	1	0	1	1	0	1	0	0	1	1
Jiangsu	1	0	3	3	1	0	1	1	1	2	3	1	1	1	4	3	1	1	0	1	3	2	0	2	0	0	0	3
Jianxi	1	0	3	3	1	0	1	1	1	1	3	1	1	1	3	3	1	1	0	1	3	1	0	1	0	0	0	3
Jilin	1	1	2	1	2	0	1	0	1	0	1	3	1	1	1	1	4	2	1	1	2	3	1	3	1	1	1	1
Liaoning	3	1	2	1	4	2	3	1	0	0	1	1	0	3	1	1	2	4	2	1	2	1	1	1	1	1	2	1
Ningxia	1	1	1	0	2	1	1	1	1	0	0	0	0	1	0	0	1	2	4	2	1	0	1	0	1	1	4	0
Qinghai	0	2	2	1	1	0	0	1	0	0	1	0	1	0	1	1	1	1	2	4	2	0	2	0	2	1	2	1
Shaanxi	1	1	4	3	2	0	1	1	1	1	3	1	1	1	3	3	2	2	1	2	4	1	1	1	1	1	1	3
Shandong	1	0	1	1	1	0	1	0	1	1	1	3	1	1	2	1	3	1	0	0	1	4	0	4	0	0	0	1
Shanghai	0	3	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	2	1	0	4	0	3	1	1	0
Shanxi	1	0	1	1	1	0	1	0	1	1	1	3	1	1	2	1	3	1	0	0	1	4	0	4	0	0	0	1
Tianjin	0	2	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	2	1	0	3	0	3	1	1	0
Tibet	0	2	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	1	0	1	4	1	0
Xinjiang	1	1	1	0	2	1	1	1	1	0	0	0	0	1	0	0	1	2	4	2	1	0	1	0	1	1	4	0
Zhejiang	1	0	3	3	1	0	1	1	1	1	3	1	1	1	3	3	1	1	0	1	3	1	0	1	0	0	0	3
	34	20	47	35	43	17	31	24	19	16	35	29	17	32	39	35	38	43	27	29	47	30	20	30	18	16	27	35

d) EPA : Government Level