Forecasting Model of Structural Change in Japan Using Markov Chain

Hiroshi Sakamoto
The International Centre for the Study of East Asian Development (ICSEAD)

Working Paper Series Vol. 2012-01
January 2012

The International Centre for the Study of East Asian Development, Kitakyushu

Forecasting Model of Structural Change in Japan Using Markov Chain

Hiroshi Sakamoto†

ABSTRACT: This paper suggests the use of a forecasting model that utilizes the Markov chain when conditions show that there is insufficient time-series data. The Markov chain corresponds to the vector auto-regressive (VAR) model of the first order. If the transition probability matrix can be appropriately estimated, the forecasting model using the Markov chain can still be constructed with less time-series data even though the amount of data is sufficient to estimate the VAR model. The forecasting model, used at two points in this paper, estimates the transition probability matrix based on industrial macro data in Japan. The use of this application at a regional level will also be discussed along with changes that could possibly affect the future of industrial structure in Japan.

KEYWORDS: forecasting model, Markov chain, industrial structural change

†Research Associate Professor
The International Centre for the Study of East Asian Development, Kitakyushu (ICSEAD)
11-4 Otemachi, Kokurakita, Kitakyushu 803-0814, JAPAN
Tel: +81 93 583 6202; Fax: +81 93 583 4602   E-mail address: sakamoto@icsead.or.jp

INTRODUCTION

This paper suggests the use of a forecasting model when conditions show that there is insufficient time-series data. Even though the time series or forecasting models are variously developed, each has their merits and weak points. Therefore, it is necessary to employ the model when utilizing data only after the merits and weaknesses are realized.

In regard to time-series analysis, the main concern has changed from the simultaneous-macro equation model into the vector auto-regressive (VAR) model (Sims, 1980). The VAR model primarily descends in relation to data, and an economic explanation then becomes poor. Only the estimated point within the structural macro model converted into the reduced system provides some economic meaning. In addition, although the VAR model needs a sufficient amount of time-series data, it is easier to deal with compared to the macro model, particularly when the numbers of variables are few. Moreover, this idea is different from factor analysis that finds significant statistical variables. However, because VAR is a time-series analysis, a strict data analysis of time-series characteristics (unit root and cointegration test) is necessary.

The data in this study is in regard to industry and region, particularly the GDP. Such data is in general reported as annual data with a changing estimate method. Therefore, it is possible that the time series used is insufficient. The purpose of this study is to propose the use of a forecasting model for these instances.

The Markov chain as a forecasting model is thought as a type of VAR model that utilizes previous data to forecast the future since it corresponds to the VAR model of the first order. Therefore, readers may view it is as an extension of research from a VAR model. Since the Markov chain does not need time-series data, an examination of the VAR model's time-series characteristics becomes necessary without the possibility of the error term correlation that exists among variables. Moreover, a feature allows the possibility to estimate even by the multivariable. On the other hand, it is impossible to correspond to a complex auto regression of a higher order because an appropriate estimate method of the transition probability matrix has not been established. These are the merits and weak points of the employed model.

However, the forecasting model from the data in this study shows that the Markov chain is suitable. Regarding the complex higher-order auto regression in this study, it is safe to assume that the first order auto regression includes a lag of one year because annual data is used. Alternatively, if the transition probability matrix can be appropriately estimated, the forecasting model can still be constructed with less time-series data for the Markov chain even though there is sufficient time-series data necessary to estimate the VAR model. Therefore, the forecasting model, used at two points in this paper, utilizes a method of estimating the transition probability matrix.

This explanation of the estimate method is in the next section. Subsequently, an application based on industrial macro data in Japan will also be introduced. In addition, the use of this application at a regional level will be discussed along with changes that could possibly affect the future of industrial structure in Japan.

2. THE MODEL

First, we must remember that the model of the Markov chain is a classic, well-known tool for the derivation of probabilistic chains (Romanovski, 1948). For each Markov transition matrix $M = (p_{ij})$ with transitional probabilities, $0 \le p_{ij} \le 1$, $\sum_{i=1} p_{ij} = 1$, the linear probabilistic chain can be derived as $p_{t+1} = M\, p_t$, $t = 0, 1, 2, \ldots$ (Sonis and Dendrinos, 2009). If we apply it, the Markov transition matrix can also be used to model the dynamics of economic growth. Let F_t be the vector comprising the GDP in the industrial sector and the period t and F_{t+1} is the same for the period $t+1$. Suppose that M_t is the matrix that maps F_t onto F_{t+1}, therefore, we have

$$F_{t+1} = F_t \cdot M_t \,. \quad (1)$$

Assuming that the transition matrix M_t is time specific, the share vector after the s period F_{t+s} will be given by

$$F_{t+s} = F_t \cdot M_t \cdot M_{t+1} \cdots\cdots M_{t+s-1} = F_t \cdot \prod_{i=0}^{s-1} M_{t+i} \,. \qquad (2)$$

Therefore, the current level of the GDP is modeled by the Markov chain.

Second, we will introduce how to estimate the transition matrix M_t by using actual data. In this case, since M_t cannot be directly obtained from actual data, an estimation procedure will be necessary. The procedure implemented in this paper runs along the following lines:

If F_t is (3 x 1), the transition matrix M_t for time t will be (3 x 3) and it will look as follows:

$$M_t = \begin{pmatrix} a_{t,11} & a_{t,12} & a_{t,13} \\ a_{t,21} & a_{t,22} & a_{t,23} \\ a_{t,31} & a_{t,32} & a_{t,33} \end{pmatrix} \quad (3)$$

Suppose $F_t' = (b_{t,1}\ b_{t,2}\ b_{t,3})$ and $F_{t+1}' = (b_{t+1,1}\ b_{t+1,2}\ b_{t+1,3})$. According to equation (1), we have

$$b_{t+1,1} = b_{t,1}*a_{t,11} + b_{t,2}*a_{t,12} + b_{t,3}*a_{t,13} \ (4\text{-}1)$$
$$b_{t+1,2} = b_{t,1}*a_{t,21} + b_{t,2}*a_{t,22} + b_{t,3}*a_{t,23} \ (4\text{-}2)$$
$$b_{t+1,3} = b_{t,1}*a_{t,31} + b_{t,2}*a_{t,32} + b_{t,3}*a_{t,33} \ (4\text{-}3).$$

However, in this formula, the property of the Markov chain may not hold when the sum of the column of probability matrix M_t becomes equal to 1.

$$\sum_{k=1}^{3} a_{t,jk} = 1 \quad \forall j \quad (5)$$

Therefore, since we assume that the adjustment parameter will hold the property, several ideas can be considered. However we adopt the total growth rate of GDP g_t when using an adjustment parameter. g_t is simply defined by

$$g_t = \sum_{j=1}^{3} b_{t+1,j} \Big/ \sum_{j=1}^{3} b_{t,j} \quad (6)$$

Then, we modify the equations to be

$$b_{t+1,1} = g_t\,(b_{t,1}*a_{t,11} + b_{t,2}*a_{t,12} + b_{t,3}*a_{t,13}) \text{ (4'-1)}$$
$$b_{t+1,2} = g_t\,(b_{t,1}*a_{t,21} + b_{t,2}*a_{t,22} + b_{t,3}*a_{t,23}) \text{ (4'-2)}$$
$$b_{t+1,3} = g_t\,(b_{t,1}*a_{t,31} + b_{t,2}*a_{t,32} + b_{t,3}*a_{t,33}) \text{ (4'-3).}$$

However, these three restrictions are insufficient to uniquely solve the nine elements of the matrix M_t. We will need more restrictions. In this regard, we note that one trivial solution of M_t. is the identity matrix, although it is not the desired solution. However, it can provide the source of necessary restrictions. Assuming that the distribution does not greatly vary from one period to the next, the case will be where the elements of M_t. are such that the matrix will mimic the identity matrix. Using this idea and generalizing M_t. to be n x n, we can estimate the elements of M_t. based on the following minimization procedure:

$$\text{Minimize} \quad \sum_{j=1}^{n}\sum_{k=1}^{n}\left(a_{t,jk} - i_{jk}\right)^2$$

$$\text{Subject to} \quad b_{t+1,j} = g_t \cdot \sum_{k=1}^{n} b_{t,k} \cdot a_{t,jk}\,, \quad \forall j\,,$$

$$\text{and} \quad \sum_{k=1}^{n} a_{t,jk} = 1, \quad \forall j \quad (7)$$

where i_{jk} is an element of identity matrix I and g_t is the total growth rate of GDP as before mentioned ($g_t = \sum_{j=1}^{n} b_{t+1,j} \Big/ \sum_{j=1}^{n} b_{t,j}$). This minimization problem can be solved by using non-linear programming to produce a unique solution for the elements $a_{t,jk}$.

Third, we construct the transition matrix $\bar{M}$ for forecasting. Since the above estimated transition matrix M_t is time specific, we first consider the average of the elements:

$$\overline{M} = \sum_{t=1}^{s} M_t \Big/ s \quad (8)$$

If we consider the adjustment parameter (growth rate g_t) for estimating $\bar{M}$, we suggest following the modified specification:

$$\overline{M} = \sum_{t=1}^{s} g_t \cdot M_t \Big/ \left(\sum_{t=1}^{s} g_t \Big/ s \right) \quad (8')$$

or

$$\overline{M} = \sum_{t=1}^{s} g_t \cdot M_t \Big/ s \quad (8'')$$

The former keeps equation (5) while the latter does not. However, because a forecast is the goal of this paper, the model also includes the previous growth rate of the GDP. Therefore, the transition matrix that utilizes it becomes equation (8’’).

Table 1 Industry Specification

		Japan	FP	KK	FC
1. Industries					
(1) Agriculture, forestry, and fishing	a. Agriculture	a001	a001	a001	a001
	b. Forestry	a002	a002	a002	a002
	c. Fishing	a003	a003	a003	a003
(2) Mining		i004	i004	i004	i004
(3) Manufacturing	a. Food products and beverages	i005	i005	i005	i005
	b. Textiles	i006	i006	i006	
	c. Pulp, paper, and paper products	i007	i007	i007	
	d. Chemicals	i008	i008	i008	
	e. Petroleum and coal products	i009	i009		
	f. Non-metallic mineral products	i010	i010	i010	
	g. Iron and steel	i011	i011	i011	
	h. Non-ferrous metals	i012			
	i. Fabricated metal products	i013	i013	i013	
	j. Machinery	i014	i014	i014	
	k. Electrical machinery, equipment, and supplies	i015	i015	i015	
	l. Transport equipment	i016	i016	i016	
	m. Precision instruments	i017	i017	i017	
	n. Wearing apparel and clothing accessories	i018	i018	i018	
	o. Wood and of wooden products	i019			
	p. Furniture	i020			
	q. Publishing and printing	i021			
	r. Leather, fur products, and miscellaneous leather products	i022			
	s. Rubber products	i023			
	t. Others	i024			
(4) Construction		i025	i025	i025	i025
(5) Electricity, gas and water supply	a. Electricity supply	s026	s026	s026	s026
	b. Gas and water supply	s027			
(6) Wholesale and retail trade	a. Wholesale trade	s028	s028	s028	s028
	b. Retail trade	s029			
(7) Finance and insurance		s030	s030	s030	s030
(8) Real estate	a. Renting of dwellings	s031	s031	s031	s031
	b. Other real estate	s032			
(9) Transport and communications	a. Transport	s033	s033	s033	s033
	b. Communications	s034			
(10) Service activities	a. Community and social service activities	s035	s035	s035	s035
	b. Business activities	s036			
	c. Personal service activities	s037			
2. Producers of government services					
(1) Electricity, gas and water supply		s038	s038	s038	s038
(2) Service activities		s039	s039	s039	s039
(3) Public administration		s040	s040	s040	s040
3. Producers of private non-profit services to households					
(1) Education		s041	s041	s041	s041
(2) Others		s042			

(Source) Cabinet Office, Government of Japan

3. DATA

The forecast using the transition probability matrix of equation (8'') are based on industrial data in Japan. Moreover, the forecast based on industrial data from the Fukuoka Prefecture, Kitakyushu City, and Fukuoka City is reported as an example of regional data. As for the Fukuoka Prefecture, the government-designated Kitakyushu City and Fukuoka City exist. These three administrations are respectively independent even though each includes a hierarchical government structure.

The data, published on the homepage of the Cabinet Office of Japan, covers an overall period from 1990 to 2009. Data from the Fukuoka Prefecture, Kitakyushu City, and Fukuoka City are only from 1996 to 2007. Older data were not used. Therefore, an adequate number of samples were not necessarily prepared in the time-series analysis. Moreover, as summarized in Table 1, there are 42 types of industry in Japan from the Fukuoka Prefecture, Kitakyushu City, and Fukuoka City represented by FP, KK, and FC, respectively. For example, "i005" of FC represents the entire category of manufacturing. Because the substance value of "i009" of KK is negatively estimated, these are added into "i008." Therefore, we have a 42 x 42 matrix estimate in Japan and a smaller matrix for other samples. The General Algebraic Modeling System (GAMS) minimization program is used to conduct the estimation.

4. ESTIMATION RESULT

The forecast was estimated by taking each sample's starting point from the final period of the sample and continuing the Markov chain until 2020. While estimating, the average rate of growth for the sample period is reflected. Therefore, the entire economy expands based on the value of the growth rate, and the possibility of decline is included in the forecast. For example, the average growth rate of Japan is 0.73% with the FP at 0.83%, and FC at 0.70%. The average growth rate of KK showed a decline of −0.05%, possibly due to Japan's ailing economy.

The graphs in Figures 1 and 2 forecast the future of the GDP of Japan according to industry. Actual numbers are shown in Table 2. As a result, the growth and decline of the industry becomes clear. It is interesting to note that although production is shown to initially decrease, it is seen in the future as long as the graph. This service-producing industry is a result of the industrial structure because Japan is an advanced country (as seen in s030, s033, and s037). Alternatively, the industry that greatly expands its production in the future includes i015, s031, and s036.

Figures 3 to 5 show predictions for the Fukuoka Prefecture, Kitakyushu City, and Fukuoka City, respectively. Actual numbers are shown in Table 3. The industrial structure does show some differences with the weight of the service industry being high while the weight of agriculture is low. In addition, Fukuoka City's ratio of manufacturing is also low. On the other hand, some manufacturing exists in Kitakyushu City mostly due to its prosperity as a "steel town." However, the center of industry understands that these are service industries (specifically, s035) that usually exist with a tendency of growth toward the future.

Figures 6 to 9 are a summary while actual figures for 2010, 2015, and 2020 are shown in Table 4. With the assumption that four figures be comparable, industry has been decreased from 42 or less to 11. Industrial structure is still considered to be changing. Substantial change is not easily seen because only the change from one year to another is shown. Industrial structure is still lead by service industries even though some manufacturing still exists. For example, the ratio of wholesale versus retail trade (s028) is higher than manufacturing (i005) in Fukuoka City.

Figure 1 Forecasting Result in Japan (a001-i021)

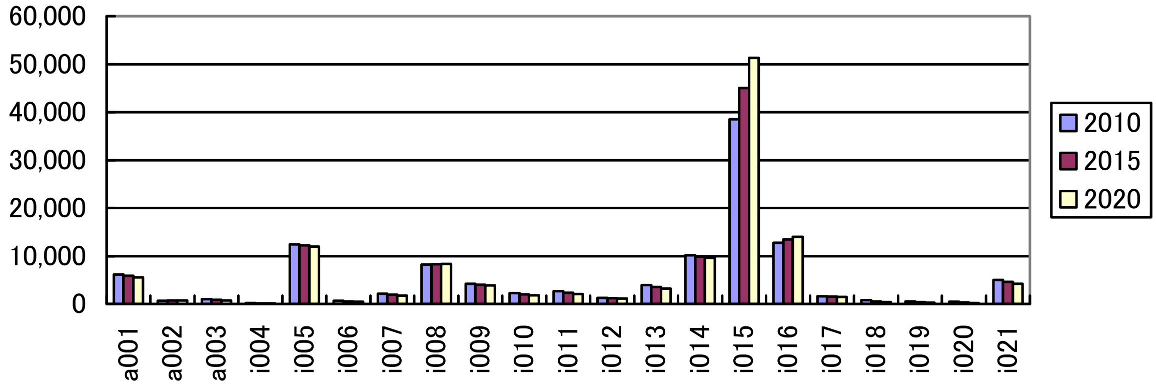

Figure 2 Forecasting Result in Japan (i022-s042)

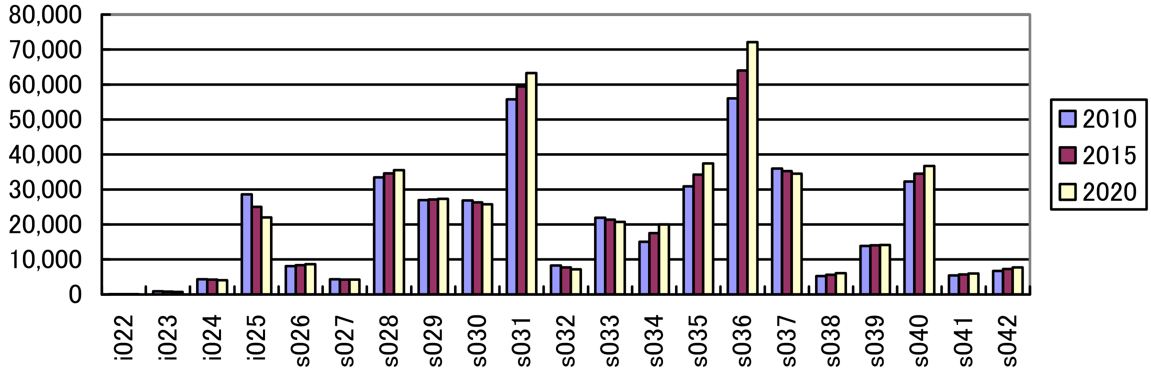

Figure 3 Forecasting Result in Fukuoka Prefecture

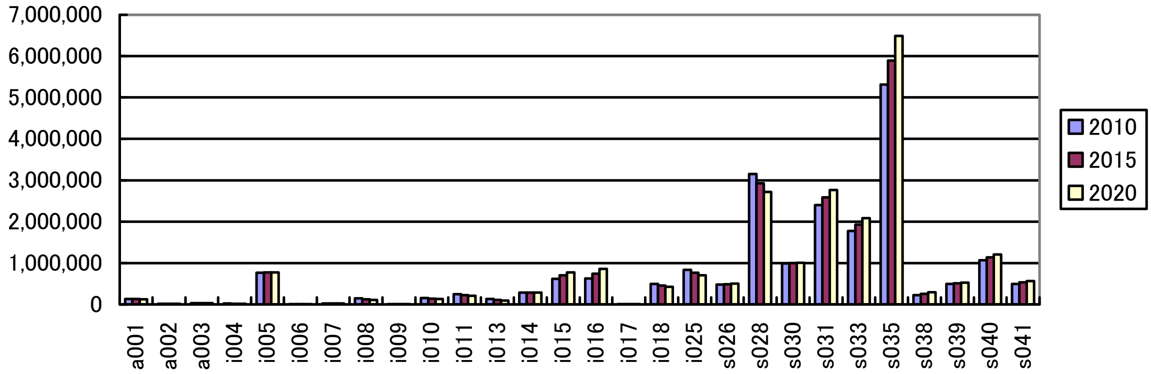

Figure 4 Forecasting Result in Kitakyushu City

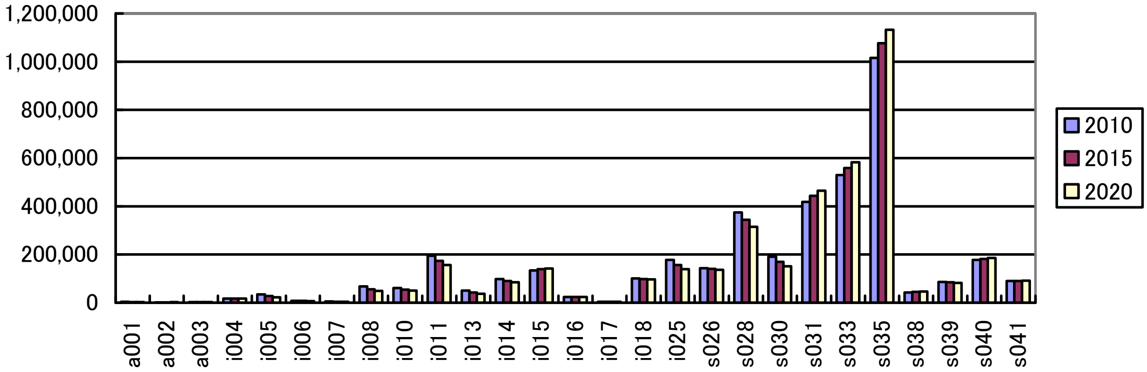

Figure 5 Forecasting Result in Fukuoka City

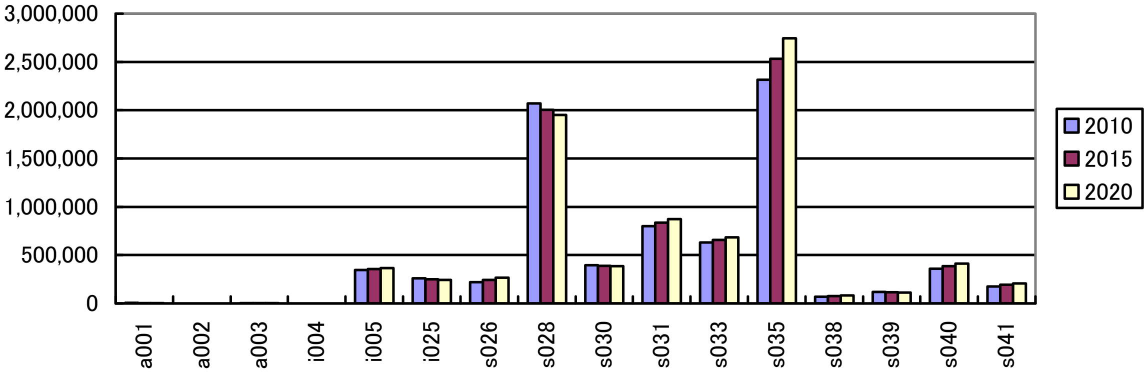

Figure 6 Changes Share in Japan

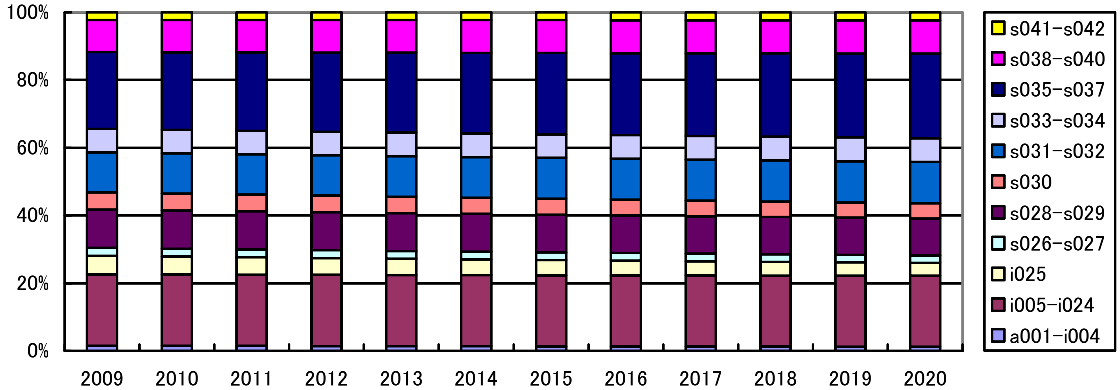

Figure 7 Changes Share in Fukuoka prefecture

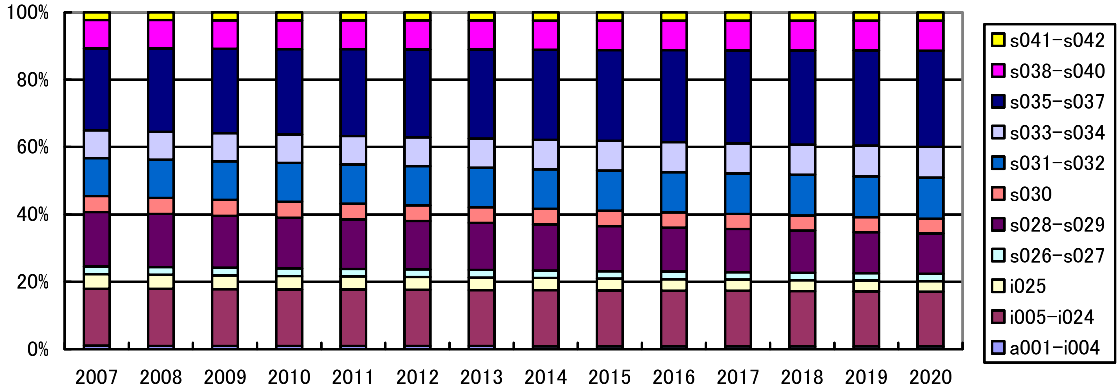

Figure 8 Changes Share in Kitakyushu City

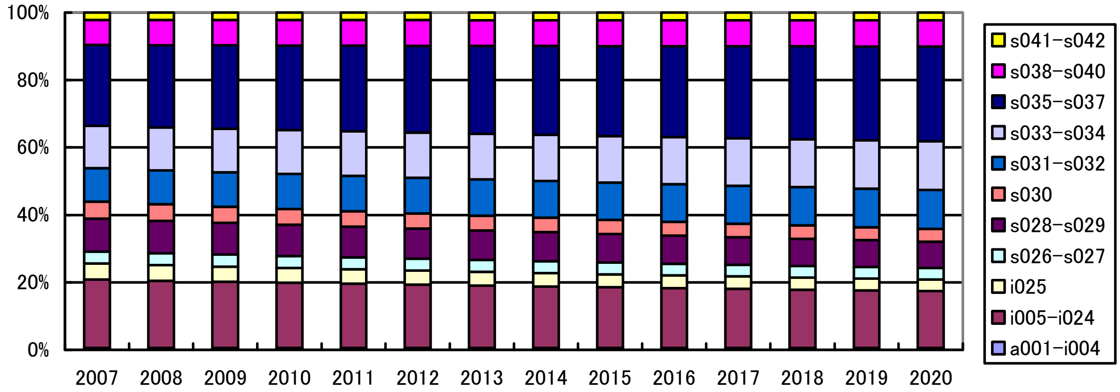

Figure 9 Changes Share in Fukuoka City

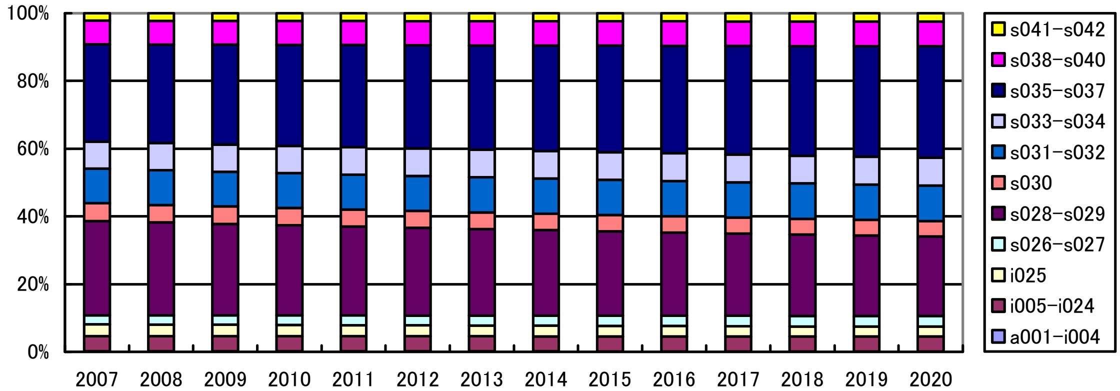

Table 2 Forecasting Result in Japan (1,000,000,000 yen)

	2010	2015	2020		2010	2015	2020
a001	6,164.6	5,852.9	5,563.4	i022	105.4	72.5	49.9
a002	689.5	720.6	742.0	i023	927.3	845.9	772.4
a003	1,010.7	847.8	716.3	i024	4,337.6	4,177.7	4,026.0
i004	208.8	160.0	126.1	i025	28,648.0	25,054.0	21,992.1
i005	12,468.8	12,213.2	11,955.3	s026	8,066.2	8,364.4	8,621.0
i006	679.2	564.5	475.5	s027	4,344.5	4,263.8	4,179.3
i007	2,149.5	1,918.4	1,719.1	s028	33,477.3	34,556.7	35,529.6
i008	8,211.8	8,295.0	8,358.4	s029	26,967.6	27,151.9	27,334.4
i009	4,208.6	4,018.8	3,847.5	s030	26,857.3	26,323.3	25,820.4
i010	2,284.5	2,035.2	1,818.9	s031	55,746.7	59,495.0	63,259.9
i011	2,695.2	2,356.1	2,083.6	s032	8,216.1	7,679.1	7,198.1
i012	1,287.3	1,199.3	1,120.4	s033	21,952.7	21,335.7	20,722.1
i013	3,952.6	3,557.5	3,204.9	s034	15,070.6	17,536.9	19,968.4
i014	10,136.8	9,889.9	9,649.7	s035	30,957.9	34,176.6	37,421.0
i015	38,525.8	45,028.1	51,305.0	s036	56,073.0	64,009.3	72,076.8
i016	12,795.5	13,446.4	14,004.5	s037	35,930.2	35,205.6	34,505.7
i017	1,622.3	1,529.8	1,445.4	s038	5,207.0	5,624.7	6,026.0
i018	792.3	537.9	372.0	s039	13,885.1	14,037.4	14,167.2
i019	515.3	387.4	291.7	s040	32,321.5	34,496.1	36,672.7
i020	447.5	316.2	225.2	s041	5,389.5	5,670.9	5,933.4
i021	5,026.0	4,606.2	4,224.4	s042	6,664.6	7,215.2	7,730.9

Table 3 Forecasting Result in Other Samples (1,000,000 yen)

	Fukuoka Prefecture			Kitakyushu City			Fukuoka City		
	2010	2015	2020	2010	2015	2020	2010	2015	2020
a001	132,970	129,460	126,077	3,367	3,072	2,804	5,261	4,904	4,597
a002	12,850	13,425	13,786	1,318	1,694	1,999	854	877	898
a003	30,948	31,059	30,739	2,074	2,123	2,106	4,791	4,546	4,338
i004	24,002	18,524	14,660	17,538	17,108	16,599	1,050	1,299	1,450
i005	763,768	770,806	776,182	34,069	27,388	22,642	346,552	355,980	364,313
i006	7,911	7,245	6,644	7,429	7,328	7,087			
i007	25,378	22,630	20,262	5,003	4,149	3,533			
i008	145,247	125,224	109,373	67,307	56,354	48,782			
i009	4,616	4,440	4,212						
i010	151,176	139,445	130,884	60,493	55,028	50,505			
i011	246,927	225,669	207,840	194,491	173,413	155,994			
i013	134,833	110,439	90,839	50,065	42,727	36,984			
i014	283,885	286,215	286,938	98,792	90,868	84,503			
i015	618,247	705,920	778,160	134,587	139,717	142,098			
i016	630,340	746,814	858,180	23,737	24,392	24,273			
i017	7,014	7,263	7,305	3,536	3,617	3,548			
i018	492,613	456,710	423,886	100,529	98,722	96,528			
i025	838,790	765,814	704,914	178,079	156,712	139,998	259,272	250,447	243,200
s026	477,858	490,583	502,492	143,598	140,211	136,445	218,616	242,619	266,456
s028	3,151,162	2,925,335	2,719,515	374,979	343,479	314,944	2,070,827	2,003,065	1,950,281
s030	994,123	1,001,077	1,006,331	190,573	169,577	151,604	394,558	389,509	384,551
s031	2,403,767	2,584,992	2,766,902	418,073	443,393	464,849	800,642	837,617	872,734
s033	1,769,795	1,925,898	2,082,054	529,084	558,222	582,187	628,872	657,790	685,229
s035	5,315,104	5,896,483	6,491,554	1,014,927	1,076,901	1,131,739	2,315,941	2,531,240	2,745,883
s038	225,445	259,079	293,835	41,827	44,590	46,708	70,158	76,941	83,489
s039	494,301	509,044	523,790	86,644	84,426	82,035	120,464	117,254	114,475
s040	1,071,218	1,141,913	1,210,883	177,540	182,335	185,787	360,034	385,626	409,978
s041	496,033	532,983	567,729	89,618	90,703	90,972	177,471	192,243	206,522

Table 4 Share Result (All Samples)

	Japan			Fukuoka Prefecture			Kitakyushu City			Fukuoka City		
	2010	2015	2020	2010	2015	2020	2010	2015	2020	2010	2015	2020
a001-i004	1.50	1.36	1.24	0.96	0.88	0.81	0.60	0.59	0.58	0.15	0.14	0.14
i005-i024	21.07	21.01	20.95	16.76	16.53	16.26	19.26	17.92	16.80	4.46	4.42	4.37
i025	5.33	4.50	3.81	4.00	3.51	3.10	4.40	3.88	3.48	3.33	3.11	2.92
s026-s027	2.31	2.27	2.22	2.28	2.25	2.21	3.55	3.47	3.39	2.81	3.01	3.20
s028-s029	11.26	11.08	10.89	15.04	13.40	11.95	9.26	8.51	7.82	26.63	24.88	23.39
s030	5.00	4.73	4.47	4.75	4.58	4.42	4.71	4.20	3.76	5.07	4.84	4.61
s031-s032	11.91	12.06	12.21	11.47	11.84	12.16	10.32	10.98	11.54	10.30	10.40	10.47
s033-s034	6.89	6.98	7.05	8.45	8.82	9.15	13.07	13.82	14.46	8.09	8.17	8.22
s035-s037	22.90	23.96	24.95	25.37	27.01	28.53	25.06	26.67	28.10	29.79	31.44	32.93
s038-s040	9.57	9.73	9.85	8.55	8.75	8.91	7.56	7.71	7.81	7.08	7.20	7.29
s041-s042	2.24	2.31	2.37	2.37	2.44	2.49	2.21	2.25	2.26	2.28	2.39	2.48

5. CONCLUDING REMARKS

This study constructed a VAR forecasting model by using the Markov chain and an insufficient amount of time-series data. It introduced the issue by using industrial data in Japan. This method of estimating the Markov chain by solving the minimization problem might have some practicality. Moreover, the use of industrial data is not required if regional data can be utilized. However, special software is necessary to solve the problem although the use of it has yet to become widespread. Moreover, the industry has been biased toward the service industry even though Japan is known as an advanced country. In addition, this type of data should be calculated for use in developing countries.

REFERENCES

Romanovski, V.I., 1948. *Discrete Markov Chains*, Gostechizdat, Moscow. (Book)

Sims, C., 1980. Macroeconomics and Reality, *Econometrica*, 48, pp. 1–48. (Journal Articles)

Sonis, M. and D.S. Dendrinos, 2009. Socio-spatial dynamics and discrete non-linear probabilistic chains, M. Sonis and G.J.D. Hewings (eds.), *Tool Kits in Regional Science*, Springer-Verlag Berlin Heidelberg, pp. 177–197. (Book Chapter)