REGIONAL AND TEMPORAL ANALYSIS OF CAR OWNERSHIP IN JAPAN

Tadashi ITOH
Research Associate
Department of Civil Engineering Systems
Kyoto University
Yoshida-Honmachi, Sakyo-ku, Kyoto,
606-01 Japan
Fax: +81-75-753-5759
E-mail: itoh@utel.kuciv.kyoto-u.ac.jp

Haruo ISHIDA
Professor
Institute of Policy and Planning Sciences
University of Tsukuba
Tennodai, Tsukuba, Ibaraki,
305 Japan
Fax: +81-298-55-3849
E-mail: ishida@shako.sk.tsukuba.ac.jp

abstract: This paper aims to grasp the regional and temporal changes in both car ownership and the effects of various socio-economic variables on it. To do so, we collected annual car ownership and socio-economic data of each prefecture in Japan from 1965 to 1993. We provide the information of the regional and temporal changes of car ownership in Japan and the way of to grasp the regional and temporal changes. And we may conclude that panel model can be a convenient tool to explain well the regional and temporal changes, and also suggest that panel model is useful in forecasting car ownership levels.

1. INTRODUCTION

In Japan, the number of cars has rapidly increased since the 1960's motorization. In 1996, Japan has more than 70 million vehicles and 80% of those (about 56 million) are passenger cars. In the process of reaching the present situation of car ownership, there were the characteristic changes in ownership for each region in Japan, and various factors influence the change in ownership. In this paper, we look back to the change in ownership from the late 1960's to the present and make clear the time-series change and regional characteristics of car ownership. By using time series data according to the administrative division of Japan, we build models which explain car ownership using socio-economic factors and we make clear the regional and temporal change of the influence factor of car ownership. Moreover, we show the possibility to estimate the future car ownership precisely.

2. METHOD TO ANALYZE CAR OWNERSHIP

The main purpose of this study is to build models to estimate and forecast car ownership by region in Japan. One of the conventional approaches to analyze car ownership is the time series approach which uses time series data of one region and explain by growth model (*ex.* Tanner 1978). The other approach uses cross-sectional data of one time point and explain by regression model (*ex.* Button 1973). However, the conventional approaches are not sufficient to explain and forecast car ownership by region. The time series approach dose not guarantee the possibility of model transferability to other regions and cross sectional approach can not consider the structural change in the future. Therefore, by assuming that the time series data of each region is panel data , we build a model for the time series change and forecast car ownership by region by using panel analysis technique (cf. Hsiao 1986).

In this study, we apply a linear regression model which easily considers the influence factors to car ownership as the panel analysis technique. Ordinary linear regression model is shown as follows :

$$y_{it} = \beta_0 + \sum_{k=1}^{K} \beta_k x_{kit} + u_{it} \quad , \quad i = 1,\cdots,N \quad , \quad t = 1,\cdots,T. \qquad (1)$$

where y : dependent variable, x : independent variable, u : error term,
β_0 : intercept, β_k : coefficient of variable, i : region, t : year.

If we focus on the difference of car ownership itself between regions and/or time, we can and/or time as follows :

$$y_{it} = \beta_{0it} + \sum_{k=1}^{K} \beta_k x_{kit} + u_{it} \quad , \quad i = 1,\cdots,N \quad , \quad t = 1,\cdots,T. \qquad (2)$$

We can also describe " Variable Coefficient Model " which consider the difference of factors' influence between regions and/or time as

$$y_{it} = \beta_{0it} + \sum_{k=1}^{K} \beta_{kit} x_{kit} + u_{it} \quad , \quad i = 1,\cdots,N \quad , \quad t = 1,\cdots,T. \qquad (3)$$

However, it is not worth to introducing dummy variables and variable coefficients of all regions and time, because a model which has many parameters cannot express the true regional and time structure. Therefore, it is necessary to grasp the characteristics of the change of car ownership as dependent variable and the regional and temporal difference of the influence of the explanation factor of car ownership.

In the following chapter, we make clear the regional and temporal characteristics of car ownership based on the change in car ownership in Japan. In chapter 4, we grasp the regional and temporal difference with the influence degrees of the factors of car ownership.

3. CHANGES OF CAR OWNERSHIP IN JAPAN

3.1 Changes in the Number of Passenger Car

Figure 1 shows the change in the number of the vehicles in Japan from 1965 to 1995. As shown in figure 1, in 1965 there are about 7.9 million vehicles in Japan, and in the 1970s the number of vehicles rapidly increased. In the end of 1995, there are about 70 million vehicles in Japan. When we focus on passenger cars and light vehicles for personal trip (we define these vehicle as passenger car , that is our target of this study) , we can find that the ratio of passenger car increases year by year. In 1965, the ratio is 44.6%, then in 1995 it reaches 77.7%.

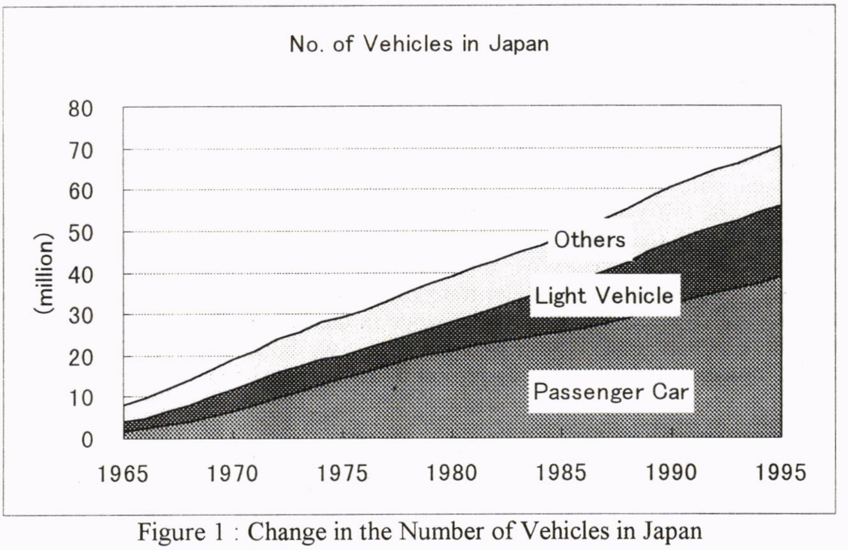

Figure 1 : Change in the Number of Vehicles in Japan

3.2 Temporal Change in Car Ownership

In order to grasp the regional and temporal change in car ownership, we look at how many cars households have, not the number of cars in the region. Therefore, we take up the number of cars per household (that is "car ownership" in this study), then we try to grasp regional and temporal characteristic of car ownership.

Temporal characteristics of car ownership are shown by the degree of increase in car ownership in a year. Figure 2 shows the change in the national average with car ownership increment in 1966-93. Based on the increment of each year, we can classify each year with socio-economic condition.

"Period 1" (1966-72) is the time of high growth of the economy and car ownership was increasing with it. In "period 2" (1973-76), the increment became low because the high growth of the economy calmed down and in 1973 the petroleum crisis happened. In "period 3" (1977-86), the petroleum crisis influence disappeared and stable increase was seen. In "period 4" (1987-90), increment became bigger because multiple car ownership in a household was very popular. In "period 5" (1991-93), with the failure of the bubble economy, the increment declines substantially.

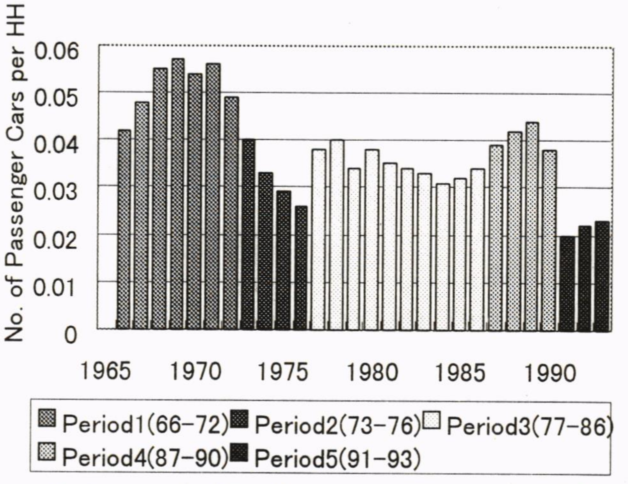

Figure 2 : Increment of Car Ownership (Japan National Average)

3.3 Change in Car Ownership for Each Prefecture

Figure 3 shows the situation of the number of passenger car per household for each prefecture of Japan in 1965, 1975, 1985, and 1995. In 1965, high values are shown in prefectures which contain big cities such as Tokyo, Aichi, and Osaka. At this time, cars were still expensive and ownership was moving ahead in the region where the levels of income were high relatively. In 1975, the spread of the car moved ahead in the whole country because of the relative decline in car prices. On the other hand, it is difficult to own cars in prefectures around Tokyo and Osaka. Because higher population density and higher land prices prevent them from having parking lots. In 1985, the ownership continued to improve favorably except around Tokyo and Osaka. However, the difference in the ownership level by prefecture was spreading gradually and ownership became high especially in the Chubu and northern Kanto areas. On the other hand, the tendency of low ownership had begun to appear in the western part of Japan. In 1995, ownership was still growing except around Tokyo and Osaka. Households that have more than two cars become very popular and in some prefectures the average number of cars per household will reach 2.0 in a few years. As for the change in ownership itself, regional characteristic is expressed in the pattern of ownership change by the difference in region such as the area around the big city and the area in the center of Honshuu.

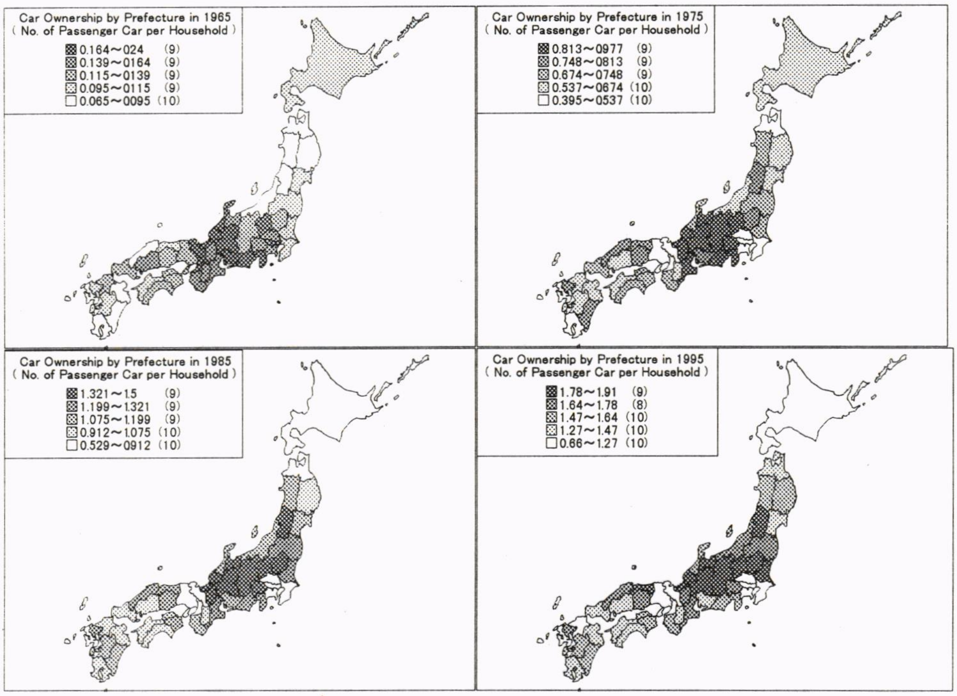

Figure 3 : Car Ownership by Prefecture

3.4 Regional Characteristics of Car Ownership

From the former analysis, we found regions where ownership was not improving though the ownership level was high in 1965 and regions where now the ownership level is high but low in 1965. That is, regional characteristics of change in car ownership are expressed in the difference of the increase tendency and the ownership level in 1965. Therefore, we classified prefectures into six regions using cluster analysis which classifies by the mean of between group of the square of Euclid distance of each prefecture plotted for the axis of the ownership level in 1965 and of the increment with the ownership level from 1965 to 93 (Figures 4 and 5).

"Region 1" is where the ownership level is always high at each time and the prefectures in Hokuriku district and so on correspond to this region. "Region 2" has a high ownership level at present with the ownership level being low at first in 1965 but improving higher within these 25 years. The prefectures in southeastern Tohoku, northern Kanto, Chogoku and Shikoku districts correspond to this. In "Region 3", the ownership level is relatively low in 1965 and the increment of the ownership is same level as the national average. The prefectures in the districts of Hokkaido, Tohoku, Shikoku and Kyushu located on both edges of Japan correspond to this. "Region 4" is the place where the ownership level became low at present after being high in 1965. Prefectures around Tokyo and Osaka and prefectures which have over a million population like Hiroshima and Fukuoka correspond to this. "Region 5" is low in the extreme development of the ownership because the ownership is restrained by the higher population density. Aichi prefecture which includes Nagoya city had a prominent ownership level in 1965, therefore we treat Aichi prefecture as the independent "Region 6".

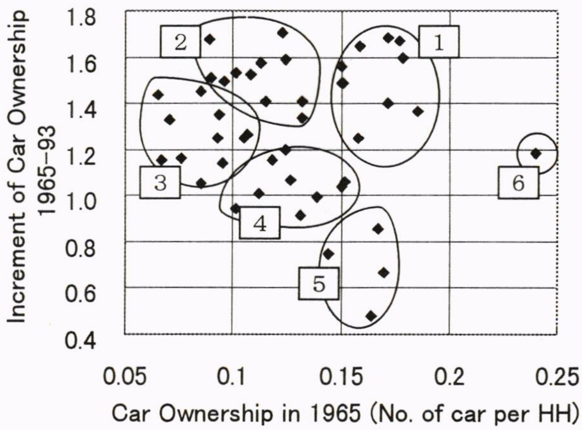

Figure 4 : Regional Characteristic of Car Ownership Change

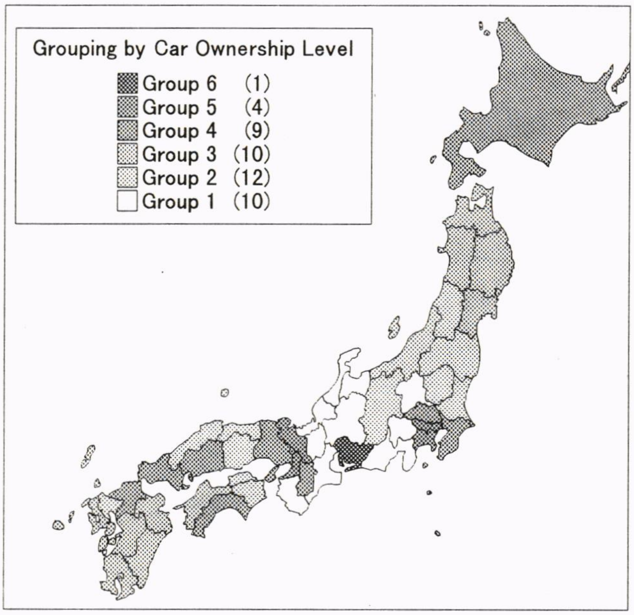

Figure 5 : Regional Grouping Based on Car Ownership Change

4. CHARACTERISTICS OF INFLUENCE FACTORS OF CAR OWNERSHIP

4.1 Selection of Factors Infuencing Car Ownership

Various socio-economic factors affect the regional differences in car ownership. In this study, we take up following three factors which affect causal relations and which are related to car usage policy.

1) Income : It is the most important factor affecting the decision for car purchase. We examine whether the relative decline of car price affects the degree of influence of income. In order to compare along time period, income variable is standardized by discount rate of

inflation rate.

2) Road Service Level : Up to now, the policy for road construction is to satisfy the demand of car usage and to develop road service levels that promote car usage. Because road construction has reached its peak and due to environmental consideration, it is now necessary to examine the present road construction policies.

3) Population Density of Densely Inhabited District (DID) : Generally, in the area with high population density, car ownership tends to be low because public transport is very convenient and it is difficult to have garage in such areas. However, in recent years, urban decentralization has spread low density areas, and the increase in car ownership is accelerating. To examine the relationship between car ownership and population density, we use the index of population density of DID which is defined as the area with more than 4,000 people per square kilometer based on the Japan National Census.

4.2 Transition of Influence Factors

We survey the change of factors in selected prefectures which have interesting characters. The selected prefectures are Tokyo, located in a metropolitan area and where car ownership is low; Gunma, located in suburb of Tokyo and where car ownership is highest; and Kagoshima, located in a local area and where the level of income is the average in Japan.

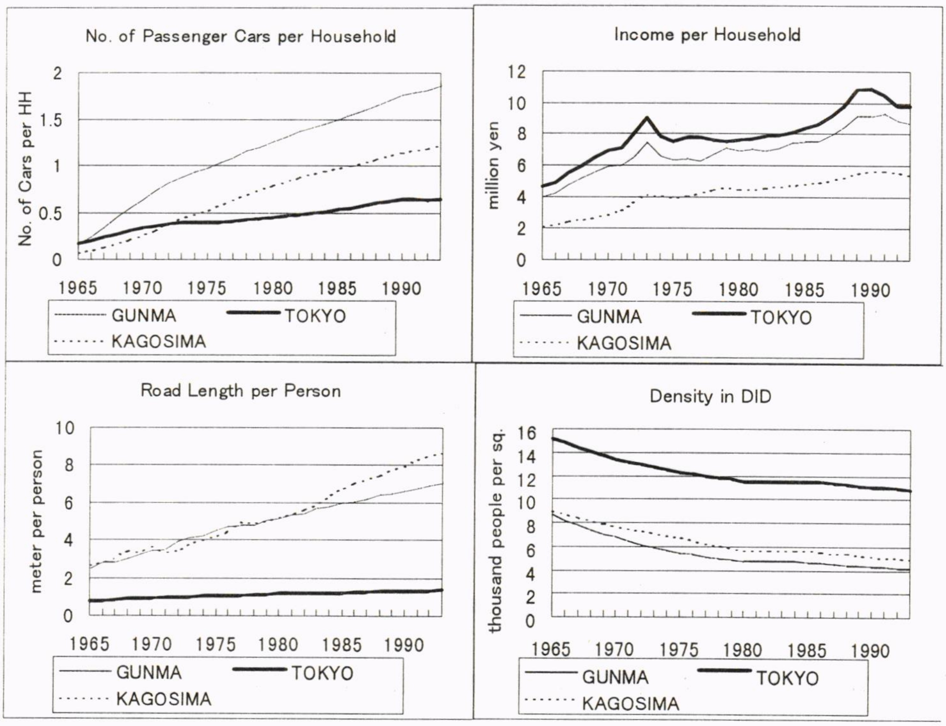

Figure 6 : Transition of Car Ownership and Influence Factors (in selected prefectures)

Figure 6 shows the transition of car ownership and three influence factors in three prefectures. In Tokyo, income per household is high, but road length per person is short and population density is high. Thus car ownership might be low. In Gunma and Kagoshima, road length per person and population density are the same level, but the difference in income appears to affect car ownership level. These are some of the typical tendencies of influence factors. In the following section, we examine all 46 prefectures and 29 time series data for 1965-93.

4.3 Temporal Characteristics of Influence Factors of Car Ownership

4.3.1 Temporal Change of Influence Factors by Simple Regression Analysis

Figure 7 shows scattergrams and simple regression lines of car ownership and each factor using data at 4 points in time (1965, 75, 85, and 93). Relation with income indicates positive influence to car ownership. The regression lines shift up year by year because slopes of the regression lines are similar, but the intercepts are getting bigger year by year. Concerning goodness-of-fit of each regression line, though the coefficient of determination (R^2) is high in 1965, more recently R^2 getting low.

In the relation with road length per person, we can find that there is no correlation in 1965 but positive correlation is strong gradually year by year. Therefore, the regression line shifts upward and R^2 gets higher, that is the relation between car ownership and road service gets stronger year by year.

Concerning the relation with DID population density, a positive relation is seen in 1965 because car ownership was growing in the big city area with high population density. However, the regression lines have negative slope in the following years and slope increases gradually. Therefore the regression line shifts upwards and relationship gets stronger year by year.

When we estimate the regression line with pooled data of every year without reference to time, we get a regression line with wrong direction. If we take consideration of the difference in slope and intercept, we can make a true estimation (Hsiao 1986).

4.3.2 Regional Difference of Influence Factors by Simple Regression Analysis

Using time series data of 29 points in time (1965-93), we estimated simple regression line of each factor by prefecture. We grasp regional difference of influence of factors by the slope of simple regression. Figure 8 shows the regional differences in regression coefficients of each factor.

Income has a positive influence on car ownership. Regions that have high income values tend to have rapidly increase in car ownership levels in the 1970's. Conversely regions that have low income values seem to be located metropolitan areas and increase in income does not tend to affect car ownership growth.

Road service also exerts a positive influence on car ownership. In metropolitan areas around Tokyo and Osaka, the value tends to be high because road length per person has not progressed within the last three decades in these areas. On the other hand, in northern Kanto and Tokai areas road construction has remarkably increased car ownership.

DID population density has a negative influence on car ownership. In metropolitan areas, as population density is high, influence of density change is not so high. However, in local or suburban areas, as population density is lower, car ownership is getting bigger. That is, low density structure is suitable for car use.

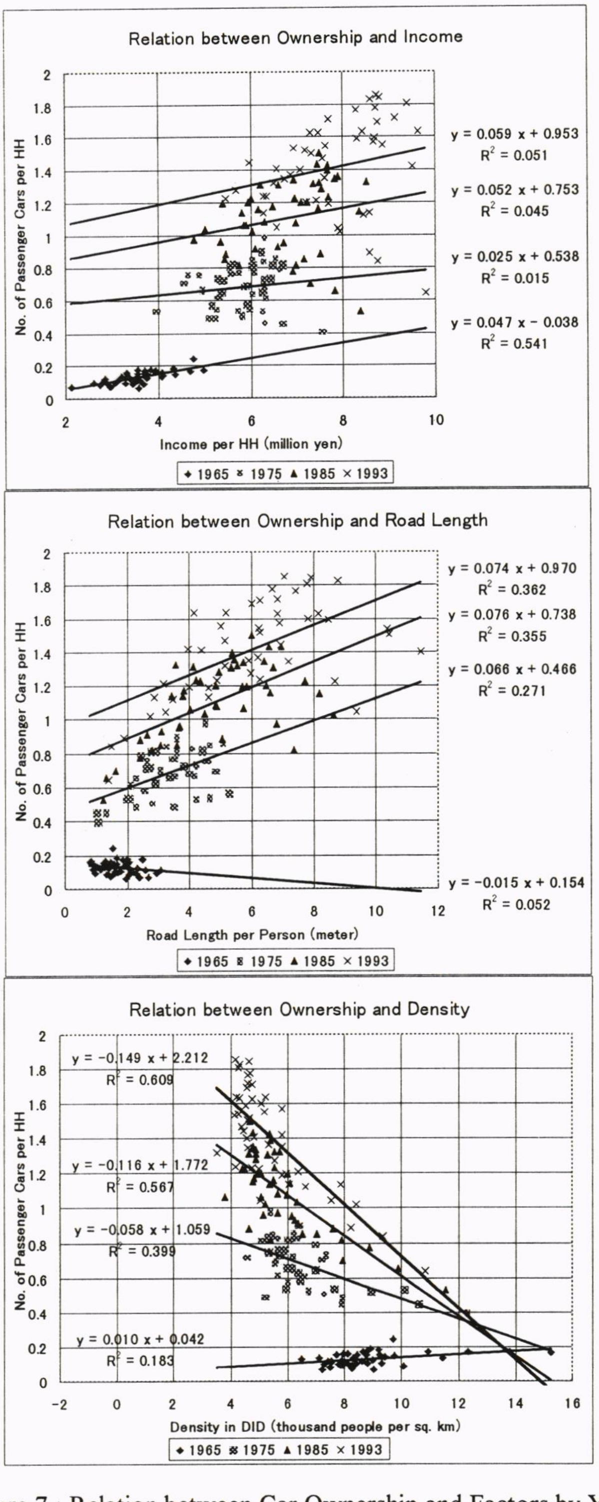

Figure 7 : Relation between Car Ownership and Factors by Year

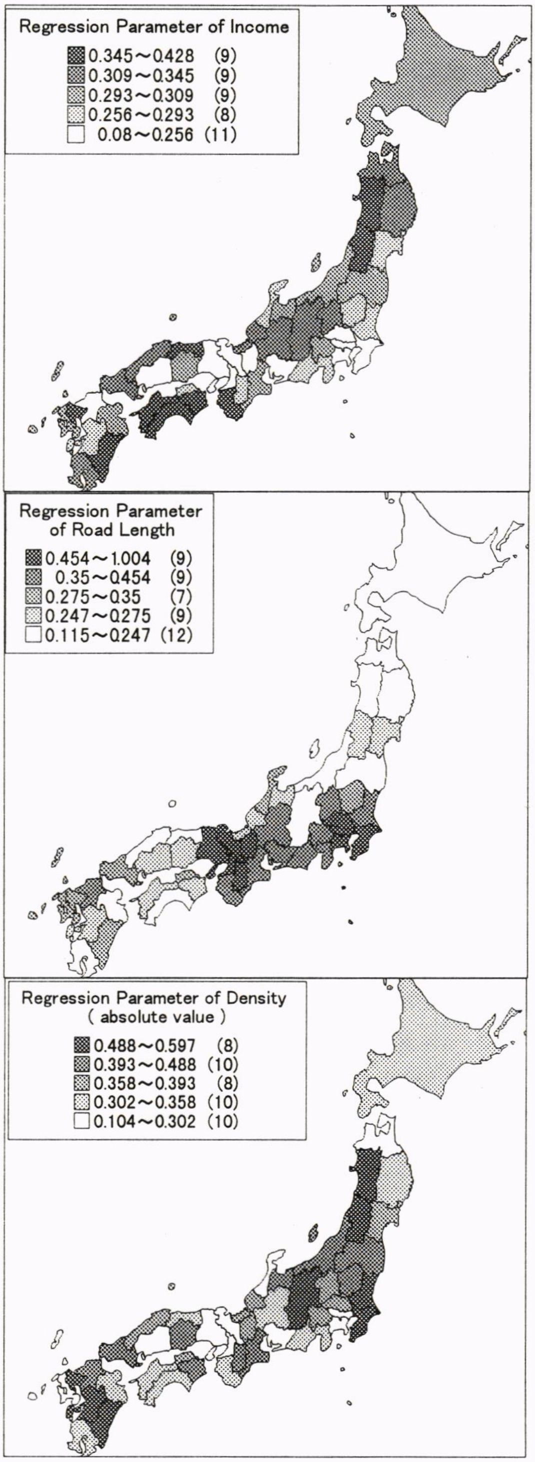

Figure 8 : Regional Differences in Influence Factors

4.4 Classification of Year and Prefecture Using Characteristics of Influence Factors

4.4.1 Classification of Year

We classify each of the 29 years based on the characteristics of temporal differences of the three influence variables. Using the three coefficients of variables of cross sectional simple regression of each time, we apply the principal component analysis to classify the years.

Calculating the eigenvalues from the correlation coefficient matrix of parameters, we reduced the principal component. 1st and 2nd components almost explain the variance of the three variables. From the relationship between components and variables, we may say that the 1st component explains "regional attributes effects" because of high correlation with road length and population density and that the 2nd component explains "income effect" because of high correlation with income (Table 1).

Table 1 : Temporal Results of Principal Component Analysis

	Correlation with Variables			Contri-bution	Interpretation of Component
	Income	Road Length	Population Density		
1st Component	0.374	0.907	-0.988	0.646	Regional Attribute Effect
2nd Component	0.925	-0.395	-0.013	0.338	Income Effect
3rd Component	0.064	0.144	0.156	0.016	

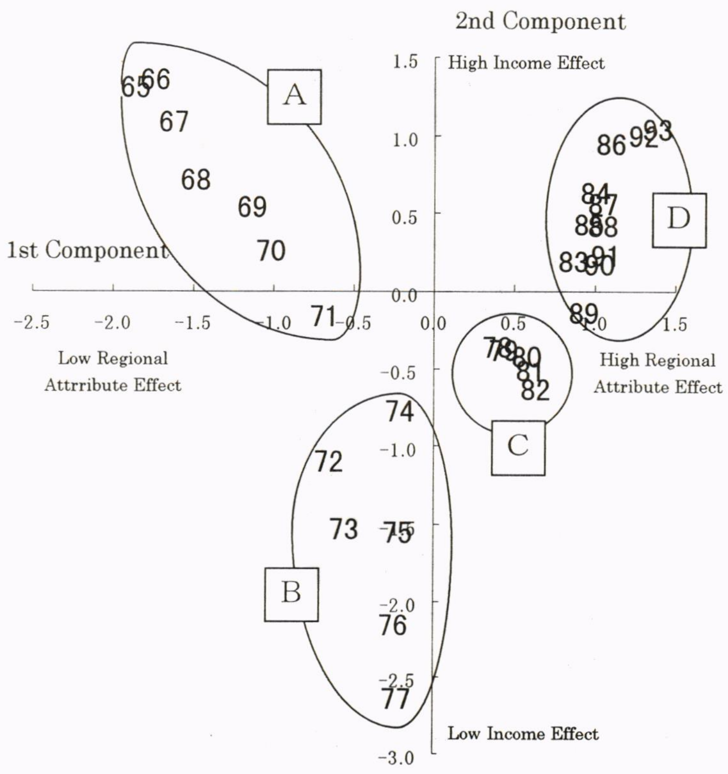

Figure 9 : Classification of Year by Principal Component Score

We made scattergram of the 1st component score versus the 2nd component score as shown in Figure 9. Then we classified time into four groups by cluster analysis. Comparing this classification with classification by car ownership increment of Figure 2, we can find relationship between increment of car ownership and the influence of factor from Table 2. For example, high growth of car ownership in Period 1 is influenced by high income effect. Stable increase of car ownership in Period 3 tends to be influenced by regional conditions such as road construction and/or density structure.

Table 2 : Comparison of Classification between Two Methods

<table>
<tr><td>Year</td><td>65</td><td>66</td><td>67</td><td>68</td><td>69</td><td>70</td><td>71</td><td>72</td><td>73</td><td>74</td><td>75</td><td>76</td><td>77</td><td>78</td><td>79</td><td>80</td><td>81</td><td>82</td><td>83</td><td>84</td><td>85</td><td>86</td><td>87</td><td>88</td><td>89</td><td>90</td><td>91</td><td>92</td><td>93</td></tr>
<tr><td>Classification by Increment</td><td colspan="8">Period 1</td><td colspan="4">2</td><td colspan="10">3</td><td colspan="4">4</td><td colspan="3">5</td></tr>
<tr><td>Tendency</td><td colspan="8">Highest</td><td colspan="4">Low</td><td colspan="10">Stable</td><td colspan="4">High</td><td colspan="3">Low</td></tr>
<tr><td>Classification by Factors</td><td colspan="7">Period A</td><td colspan="6">B</td><td colspan="5">C</td><td colspan="11">D</td></tr>
<tr><td>Attribute Effect</td><td colspan="7">Low</td><td colspan="6">Lower</td><td colspan="5">Higher</td><td colspan="11">High</td></tr>
<tr><td>Income Effect</td><td colspan="7">High</td><td colspan="6">Low</td><td colspan="5">Lower</td><td colspan="11">High</td></tr>
</table>

4.4.2 Classification of Prefecture

We classify the 46 prefectures based on characteristics of regional differences of the three influence variables. Using the three coefficients of variables of time series simple regression of each prefecture, we also apply principal component analysis as in the time classification.

Calculating the eigenvalues from the correlation coefficient matrix of parameters, we reduced the principal component. The 1st and 2nd components almost explain the variance of the three variables. From the relationship between components and variables, we may say that the 1st component explains "income effect" because of especially high correlation with income and that the 2nd component explains "regional attributes effect" because of higher correlation with road length and population density (Table 3).

We made the scattergram of the 1st component score versus the 2nd component score as shown in Figure 10. Then we classified prefectures into 8 groups by cluster analysis. Comparing this classification with the classification by car ownership difference of Figure 5, we can find the relationship between the regional difference in car ownership and the influence of factors from Figure 11. For example, if the income effect is lower, car ownership level is also lower. However, high income effect does not always promote car ownership level.

Table 3 : Regional Results of Principal Component Analysis

	Correlation with Variables			Contri-bution	Interpretation of Component
	Income	Road Length	Population Density		
1st Component	0.941	-0.848	-0.652	0.677	Income Effect
2nd Component	0.106	-0.459	0.749	0.261	Regional Attribute Effect
3rd Component	0.322	0.266	0.118	0.063	

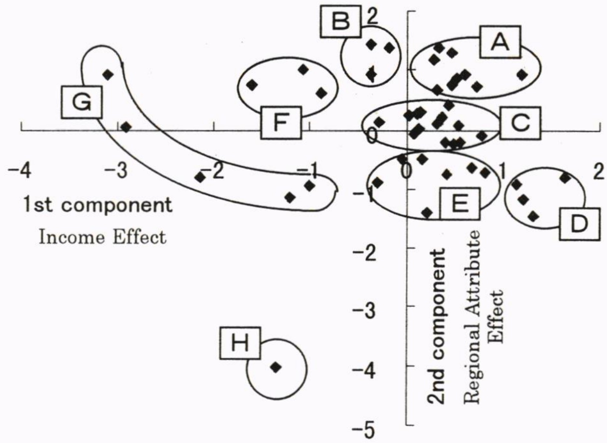

Figure 10 : Classification of Prefecture by Principal Component Score

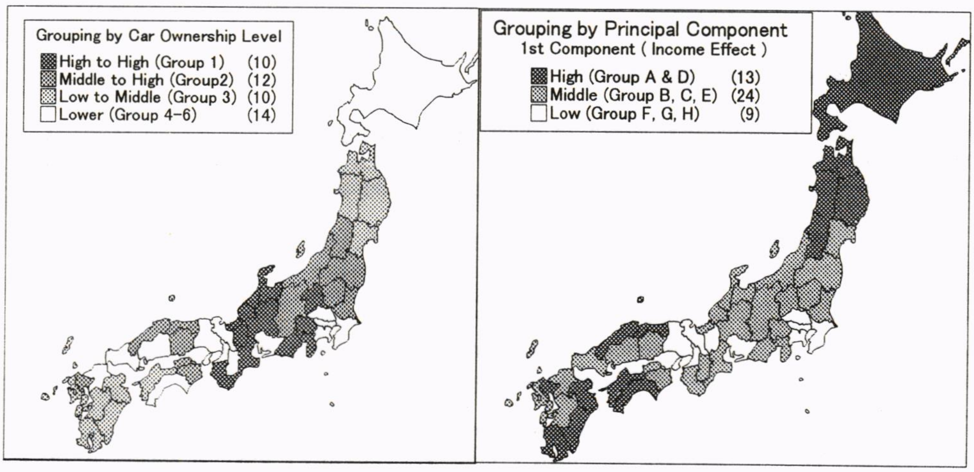

Figure 11 : Comparison between Regional Classification of Car Ownership Level and Principal Component Score

5. CAR OWNERSHIP MODELING

5.1 Models and Estimation Method

We apply panel models as “Dummy Model” and “ Variable Coefficients Model” that take regional and/or temporal difference into consideration and also apply “Ordinary Regression Model” which does not take regional and/or temporal difference into consideration.

In “Dummy Model”, to describe difference of car ownership itself, we use the 5 period classification and the 6 regional classification by car ownership level as dummy variables. In “Variable Coefficient Model”, to describe the difference of influence of factors, we apply the 4 period and the 8 region by principal component analysis as variable classification.

Therefore, we build 5 kinds of model to examine goodness of fit and precision of prediction

1) Ordinary Regression Model

$$\underset{T\times1}{\mathbf{y}_i} = \underset{T\times1}{\mathbf{e}_T}\beta_0 + \underset{T\times K}{X_i}\,\underset{K\times1}{\boldsymbol{\beta}} + \underset{T\times1}{\mathbf{u}_i}\ ,\ i=1,\cdots,N. \quad (4)$$

This is the same equation as,

$$\underset{N\times1}{\mathbf{y}_t} = \underset{N\times1}{\mathbf{e}_N}\beta_0 + \underset{N\times K}{X_t}\,\underset{K\times1}{\boldsymbol{\beta}} + \underset{N\times1}{\mathbf{u}_t}\ ,\ t=1,\cdots,T. \quad (5)$$

2) Regional Dummy Model

$$\underset{T\times1}{\mathbf{y}_i} = \underset{T\times G}{(\mathbf{e}_T\mathbf{p}_i)}\underset{G\times1}{\boldsymbol{\beta}_0} + \underset{T\times K}{X_i}\,\underset{K\times1}{\boldsymbol{\beta}} + \underset{T\times1}{\mathbf{u}_i}\ ,\ i=1,\cdots,N. \quad (6)$$

3) Temporal Dummy Model

$$\underset{N\times1}{\mathbf{y}_t} = \underset{N\times G}{(\mathbf{e}_N\mathbf{p}_t)}\underset{G\times1}{\boldsymbol{\beta}_0} + \underset{N\times K}{X_t}\,\underset{K\times1}{\boldsymbol{\beta}} + \underset{N\times1}{\mathbf{u}_t}\ ,\ t=1,\cdots,T. \quad (7)$$

4) Regional Coefficient Model

$$\underset{T\times1}{\mathbf{y}_i} = \underset{T\times(K+1)G}{\left([\mathbf{e}_T|X_i]\otimes\mathbf{p}_i\right)}\ \underset{(K+1)G\times1}{\boldsymbol{\beta}} + \underset{T\times1}{\mathbf{u}_i}\ ,\ i=1,\cdots,N. \quad (8)$$

5) Temporal Coefficient Model

$$\underset{N\times1}{\mathbf{y}_t} = \underset{N\times(K+1)G}{\left([\mathbf{e}_N|X_t]\otimes\mathbf{p}_t\right)}\ \underset{(K+1)G\times1}{\boldsymbol{\beta}} + \underset{N\times1}{\mathbf{u}_t}\ ,\ t=1,\cdots,T. \quad (9)$$

Where

y : dependent variable, X : independent variable, □ : coefficient, u : error term,
K : number of variables, G : number of group,
T : number of time, N : number of prefecture,
P : correspondent matrix (regional and prefecture ; period and year),

$$\underset{T\times1}{\mathbf{e}_T} = \begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}, \quad \text{or} \quad \underset{N\times1}{\mathbf{e}_N} = \begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix},$$

$$\underset{N\times G}{P} = \begin{bmatrix}\mathbf{p}_1\\ \mathbf{p}_2\\ \mathbf{p}_3\\ \vdots\\ \mathbf{p}_N\end{bmatrix} = \begin{bmatrix}1 & 0 & \cdots & 0\\ 1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix}, \quad \text{or} \quad \underset{T\times G}{P} = \begin{bmatrix}\mathbf{p}_1\\ \mathbf{p}_2\\ \mathbf{p}_3\\ \vdots\\ \mathbf{p}_T\end{bmatrix} = \begin{bmatrix}1 & 0 & \cdots & 0\\ 1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix},$$

$$\underset{T\times K}{X_i} = \begin{bmatrix}x_{1i1} & \cdots & x_{Ki1}\\ \vdots & & \vdots\\ x_{1iT} & \cdots & x_{KiT}\end{bmatrix}, \quad \text{or} \quad \underset{N\times K}{X_t} = \begin{bmatrix}x_{11t} & \cdots & x_{K1t}\\ \vdots & & \vdots\\ x_{1Nt} & \cdots & x_{KNt}\end{bmatrix},$$

and here we define $[\mathbf{e}_T|X_i]$ as,

$$[\mathbf{e}_T|X_i] = \begin{bmatrix}1 & x_{1i1} & \cdots & x_{Ki1}\\ \vdots & \vdots & & \vdots\\ 1 & x_{1iT} & \cdots & x_{KiT}\end{bmatrix}.$$

"Ordinary Regression Model" of equations (4) and (5) are estimated by ordinary least squares (OLS). However, in "Dummy Model" of equations (6) and (7) and "Variable Coefficient Model" of equation (8) and (9) , more than one prefecture or year are gathered into one group. Therefore, as there is individual or time series correlation between each error term, we must estimate these equations by generalized least squares (GLS). We are going to elaborate on the estimation method of Temporal Coefficient Model of equation (9).

Equation (9) can be described into one equation as,

$$\underset{NT\times 1}{\mathbf{y}} = \underset{NT\times(K+1)G}{[XP]} \cdot \underset{(K+1)G\times 1}{\boldsymbol{\beta}} + \underset{NT\times 1}{\mathbf{u}}, \quad (10)$$

where

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_T \end{bmatrix}, \ [XP] = \begin{bmatrix} [\mathbf{e}_N | X_1] \otimes \mathbf{p}_1 \\ \vdots \\ [\mathbf{e}_N | X_T] \otimes \mathbf{p}_T \end{bmatrix}, \ \mathbf{u} = \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_T \end{bmatrix}.$$

Using GLS, the estimator of coefficient $\boldsymbol{\beta}$ is,

$$\hat{\boldsymbol{\beta}} = \left([XP]'\Omega^{-1}[XP]\right)^{-1}[XP]'\Omega^{-1}\mathbf{y}, \quad (11)$$

where

$$\Omega = \Sigma \otimes I_N = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1T} \\ \vdots & & \vdots \\ \sigma_{1T} & \cdots & \sigma_{TT} \end{bmatrix} \otimes I_N, \quad (12)$$

therefore

$$\Omega^{-1} = \Sigma^{-1} \otimes I_N. \quad (13)$$

As the variance-covariance matrix Σ is unknown, we use efficient the estimator of σ (Zellner 1962) as,

$$\hat{\sigma}_{ts} = \frac{1}{N}\left(\mathbf{y}_t - ([X_t|\mathbf{e}_N] \otimes \mathbf{p}_t)\hat{\boldsymbol{\beta}}_{OLS}\right)\left(\mathbf{y}_s - ([X_s|\mathbf{e}_N] \otimes \mathbf{p}_s)\hat{\boldsymbol{\beta}}_{OLS}\right), \quad (14)$$

where

$$\hat{\boldsymbol{\beta}}_{OLS} = \left([XP]'[XP]\right)^{-1}[XP]'\mathbf{y}. \quad (15)$$

5.2 Method for Model Validation

Concerning the validation of goodness of fit of model, we use the adjusted coefficient of determination ($\bar{R}^2$), Akaike's Information Criteria (AIC), and coefficients of unequality.

$\bar{R}^2$ is used to evaluate the explanatory power of whole samples. AIC is the index that evaluates explanation of the model with the number of parameters. Suppose θ's dimension of the maximum likelihood function $L(\theta|y) = f(y|\theta)$ is p, AIC can be defined as,

$$AIC = -2\log L(\hat{\theta}|y) + 2p. \quad (16)$$

p is the penalty for increase in number of parameters. Therefore, if the value of AIC is lower, the goodness of fit is higher. In case of a linear regression model, suppose the number of samples as n and the estimate of error's covariance is $\hat{\sigma}^2$, we can calculate maximum likelihood L as,

$$L = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\hat{\sigma}^2) - \frac{n}{2} \quad (17).$$

Coefficient of inequality is the index that shows the difference between estimate and true value. For example, the coefficient of inequality U_t in time t of N samples is,

$$U_t = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_{it} - y_{it})^2}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\hat{y}_{it}^{\ 2}} + \sqrt{\frac{1}{N}\sum_{i=1}^{N}y_{it}^{\ 2}}} \quad (18).$$

That is, if U is close to 0, estimates and true values are consistent. Otherwise U close to 1 means estimates and true values are not consistent.

5.3 Results of the Regional Model

Table 4 shows the estimation result of the regional models. To easily to interpret regional characteristics, the regional groups are sorted by the order of values and the characteristics of each region are described in this Table.

Looking at the values of dummy variable parameter reflects the level of car ownership in 1965. The tendency of the variable coefficients of each region reflects the characteristics of influence factors. Then the difference in value of coefficients in each region cannot be shown in the ordinary model.

Concerning the degree of goodness of fit of whole model, $\overline{R}^2$ of two panel models are higher than ordinary models. Incorporating regional difference into a model improved model fitness. However, the AIC index shows worse in Regional Coefficient Model. From the viewpoint of number of parameters, the Regional Dummy Model is the best model for fitness.

Figure 12 shows the degree of fitness in every year. All models could not fit will in the period of the oil crisis, but the Regional Coefficient Model fitted in other periods. Regional Dummy Model has almost the same tendency as the ordinary model. However, regional panel models could not explain the time series change completely.

5.4 Results of the Temporal Model

Table 5 shows the estimation result of the temporal models. Looking at the values of parameters of the ordinary model, they are higher than those of the two panel models. It is wrong estimation in the ordinary model. Dummy parameters reflect the level of car ownership in each period. Tendency of variable coefficient of each period reflects the degree of the factors' influence. For example, income influence was high in the first period then becomes lower in the following periods. However, in recent periods the income influence becomes high again, because income is significant to own more than two cars in a household.

Concerning the degree of goodness of fit of the whole model, $\overline{R}^2$ of the two panel models are higher than for the ordinary model. Considering the temporal difference in a model is good for model fitness. However, AIC index shows worse in Temporal Coefficient Model. From the viewpoint of number of parameters, the Temporal Dummy Model is the best model for fitness.

Figure 13 shows the degree of fitness in every year. Two panel models could fit in the period of the oil crisis, and the Temporal Coefficient Model fitted better in whole period. Temporal consideration is more effective for model fitness.

6. VALIDATION OF PREDICTIVE PRECISION OF THE MODEL

6.1 Method for Validation of Predictive Precision of Model

Generally goodness of fit of model improves with the increase of the number of parameters

Table 4 : Estimated Parameters of Regional Models

	Regional Characteristics		Ordinary Regression Model		Regional Dummy Model		Regional Coefficient Model	
			Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
Intercept			0.089	2.911	–	–	–	–
	Level of Ownership							
Region 5	High to Low		–	–	0.278	4.993	–	–
1	High to High		–	–	0.248	5.937	–	–
6	High to Middle		–	–	0.212	4.142	–	–
4	Middle to Middle		–	–	0.201	4.837	–	–
2	Middle to High		–	–	0.196	5.010	–	–
3	Low to Middle		–	–	0.188	4.825	–	–
	Income Effect	Attribute Effect						
Region D	High	High	–	–	–	–	1.417	8.333
C	Middle	Middle	–	–	–	–	0.512	7.146
B	Middle	Low	–	–	–	–	0.506	6.004
A	High	Low	–	–	–	–	0.229	1.508
E	Middle	High	–	–	–	–	0.137	1.645
H	Low	High	–	–	–	–	-0.118	-0.407
F	Low	Low	–	–	–	–	-0.352	-3.980
G	Low	Middle	–	–	–	–	-0.486	-8.065
Income			0.151	57.178	0.139	39.997	–	–
Region A	High	Low	–	–	–	–	0.178	11.038
D	High	High	–	–	–	–	0.169	17.133
E	Middle	High	–	–	–	–	0.158	24.023
H	Low	High	–	–	–	–	0.131	9.094
C	Middle	Middle	–	–	–	–	0.128	15.696
G	Low	Middle	–	–	–	–	0.099	21.719
F	Low	Low	–	–	–	–	0.092	10.642
B	Middle	Low	–	–	–	–	0.082	10.455
Road			0.072	25.353	0.073	24.458	–	–
Region G	Low	Middle	–	–	–	–	0.253	19.124
F	Low	Low	–	–	–	–	0.250	13.332
H	Low	High	–	–	–	–	0.220	4.937
B	Middle	Low	–	–	–	–	0.103	15.972
E	Middle	High	–	–	–	–	0.076	17.534
C	Middle	Middle	–	–	–	–	0.072	10.422
A	High	Low	–	–	–	–	0.023	4.225
D	High	High	–	–	–	–	-0.001	-0.171
Density			-0.075	-25.033	-0.083	-21.155	–	–
Region G	Low	Middle	–	–	–	–	-0.005	-1.204
F	Low	Low	–	–	–	–	-0.019	-3.283
A	High	Low	–	–	–	–	-0.083	-5.904
B	Middle	Low	–	–	–	–	-0.084	-11.910
E	Middle	High	–	–	–	–	-0.088	-10.229
H	Low	High	–	–	–	–	-0.108	-2.876
C	Middle	Middle	–	–	–	–	-0.121	-16.991
D	High	High	–	–	–	–	-0.261	-12.711
Adjusted R2			0.897		0.899		0.915	
AIC			-1523		-1538		-1492	

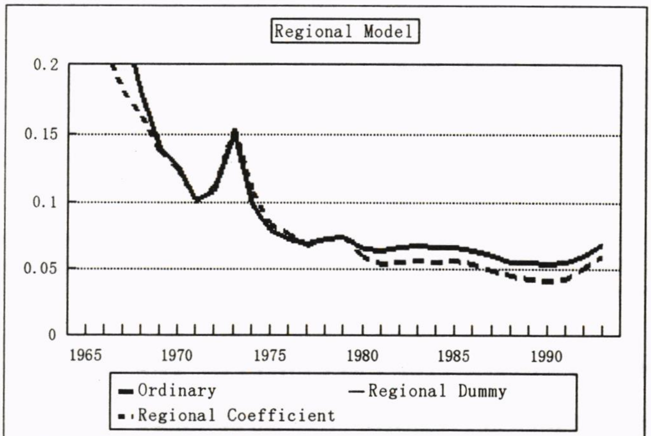

Figure 12 : Coefficient of Unequality - Regional Models -

Table 5 : Estimated Parameters of Temporal Models

	Ordinary Regression Model		Temporal Dummy Model		Temporal Coefficient Model	
	Coefficient	t-Statistic	Coefficient	t-Statistic	Coefficient	t-Statistic
Intercept	0.089	2.911	—	—	—	—
1965-72	—	—	0.335	10.659	—	—
1973-76	—	—	0.373	10.898	—	—
1977-86	—	—	0.524	15.012	—	—
1987-90	—	—	0.596	14.731	—	—
1991-93	—	—	0.668	15.969	—	—
1965-71	—	—	—	—	0.016	0.536
1972-77	—	—	—	—	0.622	13.295
1978-82	—	—	—	—	0.980	21.247
1983-93	—	—	—	—	1.062	23.662
Income	0.151	57.178	0.094	25.592	—	—
1965-71	—	—	—	—	0.099	30.075
1972-77	—	—	—	—	0.059	13.003
1978-82	—	—	—	—	0.063	13.689
1983-93	—	—	—	—	0.098	27.431
Road	0.072	25.353	0.040	13.738	—	—
1965-71	—	—	—	—	0.038	7.188
1972-77	—	—	—	—	0.036	6.826
1978-82	—	—	—	—	0.030	7.944
1983-93	—	—	—	—	0.027	8.363
Density	-0.075	-25.033	-0.065	-21.937	—	—
1965-71	—	—	—	—	-0.029	-11.831
1972-77	—	—	—	—	-0.062	-15.147
1978-82	—	—	—	—	-0.096	-21.623
1983-93	—	—	—	—	-0.125	-27.510
Adjusted R2	0.897		0.906		0.937	
AIC	-1523		-1643		-1524	

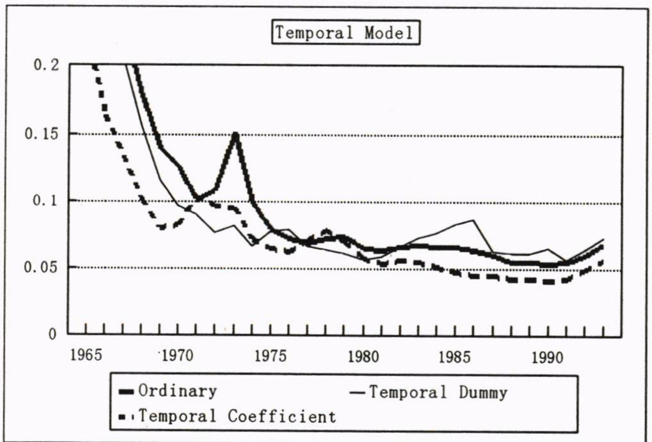

Figure 13 : Coefficient of Inequality - Temporal Models -

but predictive precision declines at the same time. In this study, models built by panel analysis technique tend to increase the number of parameters. However, panel models not only increase the number of parameters, but also present the regional and temporal structure in the model. Therefore, though the panel model has many parameters, it is suitable on a prediction model. In this chapter, we examine an extrapolation test to identify the predictive possibility of panel models.

We estimate models for different periods, one is the decade of 1965-75 and the other is the two decades of 1965-85. Then using the data of the following period, future values will be estimated by these models. To compare prediction with model structure, three kinds of models (ordinary regression, dummy model, and variable coefficient model) are estimated. Estimated models and their specifications are shown in Table 6. Future parameters of temporal models used recent parameters of the estimated period.

Table 6 : Specification of Comparative Models

Model		No. of Parameter	Adjusted R2	AIC
1965-75 Model (N = 506)	Ordinary Regression	4	0.832	-1047
	Regional Dummy	9	0.832	-1037
	Temporal Dummy	5	0.799	-953
	Regional Coefficient	32	0.867	-620
	Temporal Coefficient	8	0.854	-668
1965-85 Model (N = 1012)	Ordinary Regression	4	0.856	-1256
	Regional Dummy	9	0.858	-1268
	Temporal Dummy	6	0.862	-1294
	Regional Coefficient	32	0.881	-1242
	Temporal Coefficient	16	0.911	-1274

6.2 Comparison of Predictive Precision of the Models

To compare the degree to which the goodness of fit of model estimation and observed values are different, we calculated the coefficients of inequality for every year.

Concerning Dummy Models (Figure 14), the models of short estimated period (1965-75 Model) are rapidly worse with fitness in the prediction period. The fitness of Dummy Models in 1965-75 is worse than the ordinary model because as shown in AIC index in Table 6 these models have many parameters compared with the number of samples. On the other hand, the model of long estimated period (1965-85 Model) only the Temporal Dummy Model is worse with fitness. Increasing the number of samples improves the predictive precision.

Figure 15 shows that the Regional Model could not predict better than the ordinary model. However, the Temporal Model shows better fitness in the longer period model (1965-85 Model) and the Temporal Coefficient Model of this period has the best predictive precision (Figure 16). That is, a model which considers time series structure with more samples could have better predictive power. However, the reason why the fitness of the Temporal Dummy Model of 1965-85 is not good is a concern for future parameters. There is still room for improvement in this regard.

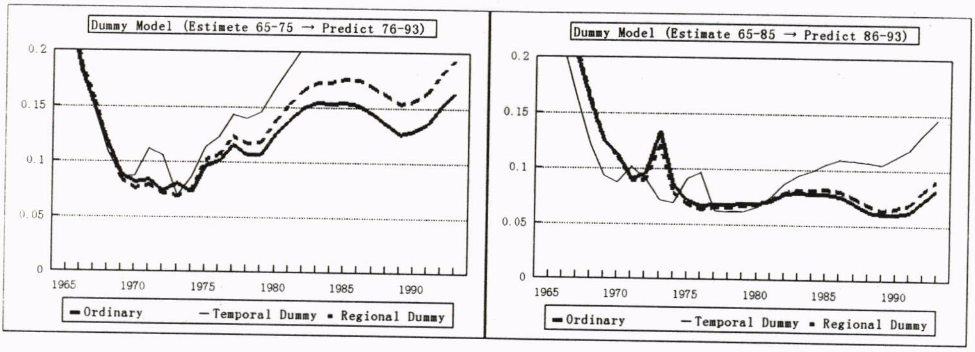

Figure 14 : Coefficient of Inequality - Dummy Models -

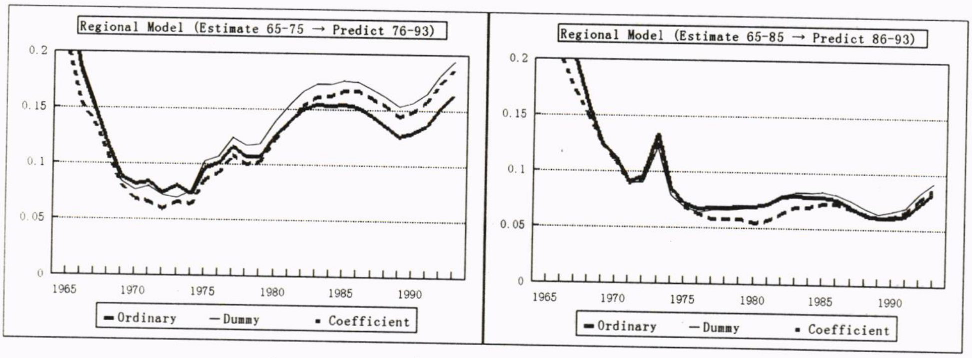

Figure 15 : Coefficient of Inequality - Regional Models -

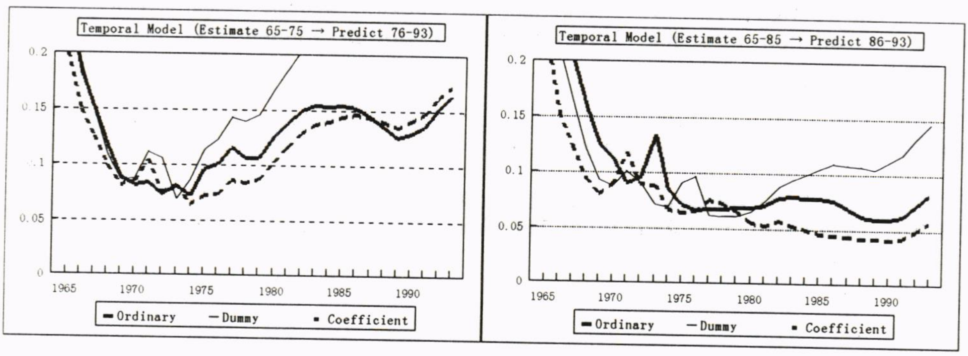

Figure 16 : Coefficient of Inequality - Temporal Models -

7. CONCLUDING REMARKS

This study developed car ownership models using panel analysis technique to grasp the regional and temporal changes in car ownership and its influence factors. To build panel models, we examined regional and temporal classification based on car ownership change itself and the factors' influence change with data on 46 prefectures from 1965 to 93 in Japan. Based on this preliminary analysis, we applied panel analysis technique and we conclude the following.

The ordinary regression models can explain car ownership well, but parameters of this model did not indicate rational values. Therefore predictive precision was not good. On the other hand, the panel model with temporal consideration shows higher level of goodness of fit , rational parameters and more precise predictive power. This is because panel models have many parameters and present temporal structure with large sample size.

For further tasks, we are trying to examine car usage in Japan. It is important to grasp the relationship among car ownership, car usage and regional attributes so as to solve problems of energy consumption and air pollution in regional level. This study is considered as a fundamental part of car usage analysis.

REFERENCES

BUTTON, K. J. (1973) Motor car ownership in the West Riding of Yorkshire : some findings. **Traffic Engineering and Control 15**, 76-78.

HSIAO, C. (1986) **Analysis of Panel Data**. Cambridge University Press.

ITOH, T. and ISHIDA, H. (1993) Panel analysis of car ownership in prefectural level. **Infrastructure Planning Review 11**, 73-79. (in Japanese)

ITOH, T. and ISHIDA, H. (1996) Regional and temporal analysis of car usage by car gasoline consumption modeling. **Infrastructure Planning Review 13**, 525-533. (in Japanese)

TANNER, J. C. (1978) Long-term forecasting of vehicle ownership and road traffic. **Journal of the Royal Statistical Society A 141**, 14-41.

ZELLNER, A. (1962) An efficient Method of estimating seemingly unrelated regressions and tests for aggregation bias. **Journal of American Statistical Association 57**, 348-368.