BAYESIAN VARIABLE AND MODEL SELECTION IN ECONOMICS: AN APPLICATION IN URBAN ECONOMICS

By

JUI-CHU TIEN

Bachelor of Business Administration
Tunghai University
Taichung, Taiwan
1988

Master of Business Administration
Louisiana Tech University
Ruston, Louisiana
1993

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 2001

BAYESIAN VARIABLE AND MODEL SELECTION

IN ECONOMICS: AN APPLICATION

IN URBAN ECONOMICS

Thesis Approved:

Thesis Advisor

Dean of the Graduate College

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to my major advisor, Dr. Lee Adkins for his intelligent supervision, constructive guidance, inspiration and friendship. My sincere appreciation extends to my other members Dr. Ron Moomaw, Dr. Dan Rickman and Dr. Tim Krehbiel, whose guidance, assistance, encouragement, and friendship are also invaluable. I would like to thank Dr. Keith Willett, Mrs. Ruby Ladd and the Department of Economics and Legal Studies for providing me with this research opportunity.

I would also like to give my special appreciation to my parents for their support and encouragement. Thanks also go to my wife, Ihsin for her support and encouragement.

Finally, I would like to thank the Department of Economics and Legal Studies for supporting me during this period of study.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

Table Page

Table Page

CHAPTER ONE

INTRODUCTION

1.1 Importance of Variable and Model Selection in Economics

Economics is a science that tries to explain and describe the characteristics and behavior of an economy (Spencer, 1971). In recent experience, certain Asian economies suffered significant contractions (especially Indonesian economy of 1999): What explains the decline in economic activity? Why did the recession in the U.S. during the 1920s and 30s become so severe? These examples deal with the whole economy (like a country), but economics also attempts to explain behavior and characteristics of smaller groups and for individuals. For example, many factors can be considered in analyzing city unemployment rates and economic theory is surely useful in narrowing the appropriate variable for consideration. For an individual, product demands, saving rates, earnings, and other outcomes of utility maximization are also explained by economic theory. But the problem is how to empirically estimate important models generated by economic theory which in turn help us to explain important features of human behavior.

In fact, there are many models that try to explain special relationships between different economic variables, such as the famous Philip's curve which states that there is an inverse relationship between the unemployment rate and changes in the money wage. That is, the higher the unemployment rate, the lower the inflation rate (Dornbusch and Fischer, 1994, p. 215). Klein (1950) constructed an empirical model of the U.S.

Macroeconomy using several equations, the parameters of which are estimated simultaneously. He includes equations for consumption, investment, private wages, equilibrium demand, private profits and capital stocks, all of which are jointly determined within the system. Other exogenous variables are included in the model, like government non-wage spending, indirect business taxes, net export, and a time trend (Green, 1993, pp. 582). These are two well-known examples of empirical economic models.

A model consists of a near representation of the system under study and is used to help us interpret, predict and make decisions. So building a model to explain economic phenomena has become very important in the discipline of economics. A model is defined as a formal or informal framework of analysis that seeks to abstract from the complexities of the real world those characteristics of an economic system which are crucial for an understanding of the behavior and the institutional and technical relationships which underlie that system. The intention is to facilitate the explanation of economic phenomena and to generate economic forecasts (Perce, 1992, p. 281).

So a main goal of economic model building is to explain the phenomena relating to the economy, and is a primary concern of economists. Before building an empirical model, it is critical to determine the variables that will be included. For example, correct specification of the Philips curve depends on expectations; there are many potential variables that affect expectations and economic theory is particularly vague on what these might be. The empirical researcher is left to use his of her own judgment in how to model this important feature of the model. If the variables we select are correct or appropriate, then we are more likely to end up with a model that can be useful for policy analysis and prediction. If a model includes irrelevant variables or omits important ones, then we end up with a poor model which is imprecise at best and misleading at worst. So, selecting appropriate variables is a critical step in building a good model in economics.

After selecting variables, there are still many possible combinations for those variables. Each combination could be deemed a model, but because a poor model will make the inferences based on the data unreliable (Burnham & Anderson, 1998), we need to be cautious when we use the model to explain the economic phenomena, especially if policy decisions are being made based on their outcome.

1.2 The Traditional Approach

Several criteria are used in variable and model selection problems. The R^2 criteria and adjusted R^2 are two of the most popular criteria used. These criteria measure the proportion of total variance accounted for by the linear influence of the explanatory variables (Judge, Griffths, Hill, Lutkepohl, & Lee, 1985, p. 862). The main disadvantage for the R^2 criteria is that it can always be made larger by adding variables to the model. Neither R^2 or adjusted R^2 account for the statistical loss which is associated with using an incorrect model.

Other classical approaches are Mallow's Cp, Amemiya's PC, Akaike Information Criterion (AIC) and the Bayesian Information Criteria (BIC). Most of the traditional model selection rules penalize the addition of regressors while rewarding corresponding the improvement in model fit. The winner is the model that represents the best result of this tradeoff under the defined norm.

Each procedure of these model selection rules, with the exception of adjusted R-square, is drawn up under a specific norm. For instance, the Cp, considers the conditional mean square predictive error, uses an estimator of the unknown parameters, and selects the model having the smallest risk (which is, selecting model that minimizing Cp). The main drawback for the Cp criteria is that if the number of explanatory variables is too large then the prediction error will become large (Judge et al., 1985). Ameniya (1980) developed a method, the PC criterion, based on the mean squared predictive error that considers the loss associated with choosing the incorrect model.

The various *information criteria* for model selection seek to balance the accuracy of the estimation and the best approximation to reality (Judge et al., 1985, p. 870). The AIC is one of the most widely used information criteria. Simulations have shown that AIC tends to choose models that are too large. The final method discussed here is the BIC, which was developed by Sawa (1978) and uses so-called pseudo-true parameter values to measure the distance between the pseudo-true parameters and the postulated parametric model. As opposed to the other classical approaches, BIC works reasonably well when the number of including variables is large. There are many well-known problems associated with using model selection rules (see Judge et al. 1995, pp 888-889). In particular, none of these procedures assures the user that the model that estimates the parameters most accurately has been chosen.

1.3 The Bayesian Approach for Variable and Model Selection

The Bayesian approach to model selection can be traced back to the 1970s. Schwarz (1978) used a Bayesian approach to derive a model selection rule similar to in spirit to the ones described above. Schwarz (1978) assumes that the posterior probability of a true model is known, and that the prior conditions of the parameters given K (number of explanatory variables) is the true model; he then uses with maximum likelihood to determine the variables to include in the design matrix (Judget et al., 1985). The Schwarz criteria incorporates prior information into analysis and provides motivations for later Bayesian approach of model selection. But the Schwarz criteria do not really consider a parameter's prior distribution as a critical element and relies on the Bayesian model asympotics. The Bayesian model selection rules used later do not suffer from these limitations.

Mitchell and Beauchmp (1988) have used a Bayesian approach for variable selection in regression analysis. Mitchell et al. (1988) assigned prior probability for each parameter (including the error term) and used usual regression equations to predict responses of dependent variables. For the independent variables, the prior probability of each variable is included or not is set as a combination of 0 or non-zero constant which is less than 1. That is, an additional parameter is used to index each coefficient in the model. The main problem is that this index parameter does not have any distribution associated with it. So we cannot know its properties or its exact relationship with other parameters. But this paper provides a good motivation for the later development of the Bayesian variable and model selection approach.

Recent developments have introduced more efficient sampling methodologies into this variable selection problem. George & McCulloch (1993, 1995, and 1997) developed a method called "stochastic search variable selection" (SSVS) to search for better subsets (or models) using a hierarchical mixture regression model. This mixture setting is similar to that of Mitchell et al. (1988) who use a spike and slab mixture. Through the development of SSVS, George and McCulloch's main innovation was to give a probability distribution to the index parameter, where they referred to a 'latent variable'.

This allows us to derive much important information from this particular distribution. George et al. (1993, 1995 &1997) also set up the prior distribution (in an objective way) for all parameters, the first step of the SSVS. The second step is to use "Gibbs sampling" which was developed by Geman and Geman (1984) to obtain the posterior distribution of the parameters. This sampling method is one of the Markov Chain Monte Carlo (MCMC) simulation techniques. George et al. used Gibbs sampling to generate a sequence of the index parameter that converges to the desired posterior distribution (according to the theory of MCMC) and contains the relevant information about the variable selected. And, this SSVS searches the promising subsets rather than evaluate the entire posterior distribution. That is the main advantage over the traditional Bayesian approach. It substantially reduces the time to evaluate the entire posterior distribution, so it is more efficient. But the SSVS approach also has one main disadvantage: the prior information is not subject to a problem-specified prior. So, an alternative approach by Brown, Vannucci and Fearn(1998) will also be considered. For this approach, Brown et al. (1998) extend George et al. (1993,1995 and 1997) to multivariate Bayesian variable selection and consider different prior settings due to the multivariate nature for the problem. This MBVS (Multivariate Bayesian Variable Selection) will be discussed in Chapter Two and will be implemented in Chapter Four.

The other Bayesian variable selection procedure is one developed by Geweke (1994). The main difference between Geweke's method and that of George et al. (1993,1995 & 1997) is the treatment of prior information. The Geweke procedure incorporates a subjective prior which could be based on the expert experience of the investigator. Like Mitchell et al. (1988), Geweke sets up a mixture of point mass at 0, but permits truncation of the parameters to a specified interval (truncated normally). So even if the considered parameter is rejected by the model selection procedure, it can still be given weight in the resulting posterior distribution. Also Geweke uses the Bayes factor in the variable selection stage by using it to compute the conditional posterior probability to indicate the parameters (detail discussed later). For computation, Gibbs sampling is implemented (similar to SSVS). The feature of Geweke's approach is that it can compute the subsets' (or model's) posterior probability for all possible subsets. But as Geweke states, the degree of collinearity must be considered because it will affect the independence of the regressors and the rate of convergence of the Markov chain.

The above three can be used exclusively for variable selection; another problem is that of to picking an appropriate model from many potential models. George (1995) evaluates the posterior probability via the Bayes factors and prior ratios of the model, but the prior is not very easy to set up. The Geweke procedure yields the posterior distribution of the parameters and the mean of this can easily be used as an estimator of the model. Still another possibility is Bayesian Model Averaging (BMA) (Raftery, Madigan and Hoeting, 1997; Volinsky, Madigan, Raftery, and Kronmal, 1997; Hoeting, Madigan, Raftery and Volinsky, 1998). Like the SSVS and the Geweke's approach, BMA also requires the posterior probability for each parameter and proper prior information for each. But BMA goes further; after computing the posterior probability of the parameters, it combines the likelihood probability and posterior probability (actually what is weighted by the posterior probability) to produce a model(s). The main advantage of the BMA is that it can account for the uncertainty of the model and the interests of the researchers (Raftery et al., 1997). Raftery et al. (1997) and Hoeting et al.

(1998) adopt two algorithms for the BMA. The first is Occam's window. This algorithm is based on the use of the Bayes factor or posterior odds ratios. If a larger model is rejected, then all nested smaller models are also rejected. Of course, we need to set up the neighborhood for the rejection (or acceptance) region to make a decision. But Occam's window has one disadvantage: the model may become inconclusive. That is, we may not be enough evidence to reject it or accept it. So this drawback of Occam's window needs to be addressed.

The second method is the MCMC approach of Madigan and York (1995) and is called the MCMC model composition (MC3) methodology. MC3 generates a stochastic sequence that translates through (or moves through) the model space. By simulating the Markov Chain many times and under certain conditions (or MCMC theorems), the average of this sequence will converge to the posterior mean for the models (Raftery et al., 1997; Hoeting et al., 1998). This BMA approach can identify proper models from a set that contains information about the model selection and can reduce to an even smaller set for more efficient computation. This will reduce the time to compute the integration of the posterior probability and the marginal likelihood for the model. Also Raftery et al. (1997) and Hoeting et al. (1998) argue that BMA has better predictive performance than other methods. So for model selection, it is argued that BMA is a good method of obtaining a suitable empirical model. Another recent paper which discusses the BMA is Raftery and Kronmal (1997). It is similar to the above papers (Hoeting et al. 1998 & Raftery et al. 1997) except that they only use the BIC as the indicator for the best model associated with the highest posterior probability.

There are several other papers that discuss or use Bayesian variables and model selection. Kuo and Mallick (1994) propose a simple approach to selecting variables using an indicator vector that is computed using MCMC. Clyde and Parmigiani (1998) use BVS (Bayesian Variable Selection) in medical studies. George and Foster (1997) use Empirical Bayes Criterion (EBC) in BVS and argue that EBC is asymptotically consistent. Clyde, Desimone and Parmigiani (1996) apply orthogronalized model mixing to the BVS. Smith and Kohn (1996) use BVS in their nonparametric regression model. Carlin and Chib (1995), Green (1995), and Dieblot and Robert (1994) emphasize the importance of MCMC in BVS. Richard (1995) discusses the PIC (F) criterion related to Bayesian model selection and does some empirical application. Chipman (1996) uses the SSVS approach but adds in the dummy variables and assumes the predictors (independent variables) have many qualitative levels. Moulton (1991) adopted the Bayesian approach to variable selection to determine the price index of radio services. But Moulton did not actually use the simulation method to solve the problem, using asymptotic approximation to get the posterior odds ratios instead. Another paper is that of Adkins, Moomaw, and Tien (1999), who adopt Geweke's approach to apply to urban economics and Brown, Vannucci and Fearn (1999) who adopt non-conjugate prior in multivariate regression model selection problem.

Most of the papers above are applications in the biological or medical fields. There are very few that apply the BVS or BMS (Bayesian Model Selection) to economics. One possible application of interest is the determinants of economic growth. There are many potential features that affect economic growth, e.g., labor, capital, income, and education (Barro, 1998).

Other factors that affect U.S. economic growth may be related to the relative size and location of major cities (such as New York, Los Angeles etc.). Jacobs (1984), uses a historical approach, to argue for the importance of the cities. Cities can gather more capital, industry, and educated workers, which in turn stimulates city growth and that the surrounding area (Jacobs, 1984). For the past 20 years, the suburban areas of the major cities have grown at a high rate, in contrast to the low growth rate of city centers. Steinackers (1998) argues that central areas attract more new firms than other locations, but the growth rate of these central areas still plays an important role in the national economic growth (Steinackers, 1998). Voith (1998) also argues that for the past 30 years, cities have affected suburban growth and national growth.

Although we know cities are critical to the nation, cities posses many characteristics that may not contribute to growth. It is important to know which one(s) are the primary factors that affect the economic growth. Glaeser, Scheinkman, and Shleifer (1995) examine the relationship between urban characteristics in 1960 and in 1990. In this paper, they provide many potential variables to be considered. Barro (1998) also suggests variables to consider: GDP, sex, education, politics, inflation rate, etc. Among these characteristics, education (or human capital) is the most often discussed. Glaeser, Kallal, Scheinkman, and Schliifer (1992) argue that knowledge spillovers affect economic growth. Rauch (1993) states that the effect of human capital is to externalize technology development: cities which have more high technology firms can offer higher wages to attract more highly educated workers (Rahch, 1993). Simon (1996) discusses the impact of human capital on English cities during a 100-year period

(1861-1961). Simon (1998) argues that the cities that have higher human capital have higher employment growth.

There are many other variables to consider: (1) The size and scale of the public sector, like the government's spending on major construction such as highways, water, sewage, and police protection (Glaeser et al., 1995). More public construction should stimulate the employment rate and promote the growth of the city; (2) Income inequalities. The magnitude of the gap between high incomes and low incomes may affect the economic growth of city (Glaeser et al., 1995); (3) The number and size of manufacturing firms. These could promote the employment growth (Glaeser et al., 1995); (4) Race. Taeuber and Taeuber (1965) provide an index to estimate the effect of segregation; (5) Other economic indexes. For example, per-capita income and the unemployment rate should be important in the growth of a city (Glaeser et al., 1995); (6) Regional effects. During the last 20 years, some regions have grown faster than others. These categories and others (such as age and technology) also will be considered.

There are many ways to measure the growth of cities, including employment growth (Simon, 1998) and population growth (Mills, 1990; Glaeser et al., 1995; Simon, 1996). The population growth rate should reflect the city's growth because immigration into a city indicates that the future of the city is good and workers have confidence in this city. As firms begin new projects and hire more workers, it will stimulate the economy of the city. So the growth of the population should be an appropriate indicator for the economic growth of the city.

BVS and BMS will be implemented to select a suitable subset of explanatory variables. By incorporating the prior information about likely parameters value associated with these variables, the BVS selects the appropriate variables according to the posterior probability, and then implements BMS (such as BMA) to select appropriate models(s). After applying the BVS and BMS to the city and MSA growth data sets, the results are compared to those obtained using more conventional procedures.

Chapter Two discusses the methodology of BVS and BMS. Chapter Three discusses the variables and observed units which will be used in this research. In Chapter Four BVS and BMS are implemented in the context of city growth models. They are also compared to the classical approach. Chapter Five concludes this research and discusses the future development of BVS and BMS.

CHAPTER TWO

METHODOLOGY FOR BAYESIAN VARIABLE AND MODEL SELECTION

2.1 The Bayesian Framework in Regression Analysis

One of the primary goals of Bayesian analysis is to derive the posterior probability associated with parameters of interest. This posterior probability combines the prior information (using the prior probability distribution) with information from the data. In other words, the posterior probability is the conditional probability of the unobserved quantities of interest which are given in the observed data (Gelman, Carlin, Stern and Rubin, 1998). The posterior probability can be written as $P(\theta|Y)$, where θ is the unknown parameter and Y is the observed data set. According to Bayesian theory, the posterior probability can be written as

$$P(\theta|Y) = \frac{P(\theta,Y)}{P(\theta)} \tag{2.1}$$

Since $P(\theta,Y) = P(\theta) \bullet \frac{P(\theta,Y)}{P(\theta)} = P(\theta) \bullet P(Y|\theta)$

(2.1) becomes $P(\theta|Y) = \dfrac{P(\theta)P(Y|\theta)}{P(Y)}$ (2.2)

Where P(Y) is the prior probability of the data set Y. This prior probability of the samples, P(Y), is often omitted because it does not depend on the unknown parameter θ.

So (2.2) becomes

$$P(\theta|Y) \propto P(\theta) \bullet P(Y|\theta) \tag{2.3}$$

The $P(\theta)$ is the prior probability for the parameter θ, which must be specified before the analysis. The prior information can be objective, depending on the information or expert knowledge of the researcher. The prior probability under a different approach will be discussed later. Next, the Bayesian approach to regression analysis is discussed.

Assume Y is the observed dependent variable, an n x 1vector, X is the observed independent variable, an n x (k+1) matrix, and ε is the n x 1 error vector. The basic equation is

$$Y = X\beta + \varepsilon \tag{2.4}$$

where β is the unknown parameter vector which is (k+1) x 1. If expert knowledge or useful information is not available, (which happens often), then a noninformative prior distribution can be specified. Assume ε ; $N(0,\sigma^2 I), I$ is an n x n identity matrix and each ε_i ($i = 1,2,3 \cdots\cdots n$) has common variance σ^2. Then the prior distribution for Y given β, σ^2 and X is a normal distribution as follows:

$$Y|\beta,\sigma^2, X ; N(X\beta, \sigma^2 I) \tag{2.5}$$

The prior distribution for β and $\ln \sigma^2$ is noninformative and chosen for convenience as (Gelman et al. 1998)

$$P(\beta, \sigma^2 | X) \propto \sigma^{-2} \tag{2.6}$$

That is, the prior distribution (or p.d.f.) for β and σ^2 under X is observed to be approximated by the inverse of σ^2. The details of informative prior distribution for β and σ^2 will be discussed later.

After specifying the prior information, the next step for the Bayesian approach is to obtain the posterior probability distribution for β and σ^2. First we determine the posterior probability for β, conditional on σ^2 and Y. Then, the posterior probability for σ^2, conditional on Y, is determined. The joint posterior distribution $P(\beta,\sigma^2|Y)$ is factored out as

$$P(\beta,\sigma^2|Y) = P(\beta|\sigma^2,Y)P(\sigma^2|Y) \tag{2.7}$$

To obtain (2.7), the first step is to specify the distribution for $(\sigma^2|Y)$, computing its variance and drawing σ^2 from the distribution $(\sigma^2|Y)$. After drawing σ^2, then draw β from the distribution of $(\beta|\sigma^2,Y)$. Once obtaining $P(\beta|\sigma^2,Y)$ and $P(\sigma^2|Y)$, (2.7) is easily derived. Obtaining the posterior probability distribution of parameters or other statistics of interest is the main objective of the Bayesian framework and simulations used to obtain these will be discussed later in this chapter.

2.2 The Stochastic Search for Variable Selection (SSVS) Approach

Section 2.1 illustrates the basic Bayesian framework for a normal regression model. An important question in regression analysis is, Which variables should be included or excluded as regressors? The importance of variable selection has already been described in the first chapter, but how does one go about selecting variables in Bayesian econometrics? Three Bayesian approaches to solve the variable selection problem are discussed below. The first is Stochastic Search Variable Selection (SSVS) derived by George and McCulloch (1993). The main goal of SSVS is to solve the classical problem of variable selection, but it differs from the traditional Bayesian approach to regression in several respects. The SSVS approach tries to find the more promising variables of the entire posterior distribution rather than to evaluate the entire posterior distribution, and this reduces the computational burden substantially. The search is done by the Markov Chain Monte Carlo (MCMC), which has gained widespread use in recent years. The two MCMC methods considered below are the Gibbs sampling and the Metropolis-Hastings (M-H) algorithms. The idea of SSVS is to use Gibbs (or M-H algorithm) to generate random samples from the posterior distribution; from the Gibbs samples, models with higher posterior probability can be identified. In the next section, the hierarchical model for SSVS is described.

2.2.1 The Hierarchical Model for SSVS

The statistical model used in SSVS is a regular regression model with normal mixture:

$$Y = X_1\beta_1 + X_2\beta_2 + \cdots\cdots\cdots\cdots + \varepsilon \text{ or } Y = X\beta + \varepsilon$$

$$Y|\beta,\sigma^2 \sim N(X\beta,\sigma^2 I_n)$$

which is the same as that used in (2.5) and (2.4). The vector Y is n x 1; X is n x k, β is a n x k and ε is n x 1 vector. This differs from (2.4) and (2.5) only in that the intercept term is ignored.

The VS (variable selection) problem arises because one or more elements of the unknown parameter vector β are either equal to zero or are small enough to be discarded from this model. So we use a latent variable $\gamma_i = 0$ or 1 ($i = 1,2,\cdots K$) to index each parameter. For $\gamma_i = 1$ means β_i has a large coefficient; if $\gamma_i = 0$ then it does not. The index parameter γ_i can be expressed as

$$P(\beta_i|\gamma_i) = (1-\gamma_i)N(0,\tau_i^2) + \gamma_i N(0, C_i^2\tau_i^2) \tag{2.8}$$

and $$P(\gamma_i = 1) = 1 - P(\gamma_i = 0) = \theta_i \tag{2.9}$$

and $$\gamma = (\gamma_1, \gamma_2, \cdots\cdots, \gamma_k) \tag{2.10}$$

where τ_i and C_i are hyperparameters.

The idea is to set τ_i small so that if $\gamma_i = 0$ then β_i is small enough to be considered equal to zero. Or, if $\gamma_i = 1$, then β_i is not zero. The parameter θ_i is the prior probability that β_i is not equal to zero. Because the value of γ_i is unknown, we need a prior mixture to obtain it and this is expressed as

$$P(\beta, \sigma^2, \gamma) = P(\beta|\sigma^2, \gamma)P(\sigma^2|\gamma)P(\gamma) \tag{2.11}$$

Using a multivariate normal distribution, the first term in (2.11) becomes

$$P(\beta|\sigma^2, \gamma) = N(0, D_\gamma R_\gamma D_\gamma) \tag{2.12}$$

where D_γ is a diagonal normal matrix and R_γ is the prior correlation matrix. The diagonal matrix D_γ is

$$D_\gamma = \{b_1\tau_1, b_2\tau_2, \cdots\cdots, b_i\tau_i\} \tag{2.13}$$

where $b_i = 1$ if $\gamma_i = 0$ and $b_i = C_i$ if $\gamma_i = 1$. So D_γ forms the prior covariance matrix, which is consistent with (2.8). Here it is assumed that β and σ are *a priori* independent given γ_i so that (2.12) can be obtained. The specific selection of C_i, τ_i and R_γ is discussed later.

For the prior distribution of σ^2, the usual inverse Gamma conjugate prior is used. That is,

$$P(\sigma^2|\gamma) = InvG(\upsilon/2, \upsilon\lambda_\gamma/2) \tag{2.14}$$

which is equivalent to $\upsilon\lambda_\gamma/\sigma^2 \sim \chi^2_\gamma$. The parameter σ^2 and λ_γ is considered to be the prior estimate of σ^2 and υ is the prior sample size. One possible choice as a prior value of λ_γ is the LS (least square) estimate of σ^2 from the linear regression of X on Y.

The prior for the latent variable γ should reflect the importance of the parameter β, that is whether a particular β_i should be included in the model or not. This type of prior information is based on the expert knowledge possessed about those variables. One expects that the better the expert knowledge, the better the outcome of the variable selection procedure. One simple but useful prior distribution is the independent Bernoulli:

$$P(\gamma) = \prod_{i=1}^{k} \theta_i^{\gamma_i}(1-\theta_i)^{(1-\gamma_i)} \tag{2.15}$$

A special case is $P(\gamma) = \frac{1}{2^k}$ where each variable has an equal chance of being included in the model. We could also put more weight on some variables and less on others using $P(\gamma)$ if desired.

For the above hierarchical set up, the latent variable vector $\gamma = (\gamma_1, \gamma_2, \cdots, \gamma_k)$ contains the useful information for the variable selection. If γ is known then, with appropriately chosen $\tau_1, \tau_2, \cdots, \tau_k$ and $C_1, C_2, \cdots, C_k$, a proper model could be obtained by including the variables (or X_i) for which $\gamma_i = 1$ and excluding variables for which $\gamma_i = 0$. But as stated as above, the expert knowledge required for this choice is not always available. However, the posterior distribution, $P(\gamma|Y)$, may provide useful information about variable selection; those γ_i with higher posterior probability identify models that are supported by the data (Y) and the prior information (γ). Also the posterior $P(\gamma|Y)$ updates the prior probability for each of the 2^k possible values of γ. From (2.2), the basic Bayesian framework,

$$P(\gamma|Y) = \frac{P(\gamma)P(Y|\gamma)}{P(Y)}$$ which can be approximated as

$$P(\gamma|Y) \propto P(\gamma) \bullet P(Y|\gamma) \tag{2.16}$$

which assume $P(Y)$ is a constant. Now the goal is to obtain the posterior distribution $P(\gamma|Y)$. In order to accomplish this one must select appropriate values for τ_i, C_i and R_γ that can be used in the computation of the hierarchical model.

2.2.2 Choice of τ_i, C_i and R_γ

From (2.8), the distribution for β_i under γ_i is a mixture of two normal distributions. To incorporate this hierarchical mixture set-up in the variable selection, τ_i and $C_i^2\ \tau_i^2$ are set to small and large values, respectively. So the $N(0, \tau_i^2)$ is a concentrated distribution and $N(0, C_i^2\tau_i^2)$ is a diffuse distribution. Assume δ_i is the intersection of these two distributions ($N(0, \tau_i^2)$ and $N(0, C_i^2\tau_i^2)$). The idea is that when the data support $\gamma_i = 1$, then X_i should probably be included in the model.

The region where $N(0, \tau_i^2)$ covers (or larger) $N(0, C_i^2\tau_i^2)$ corresponds to $|\beta_i| \le \delta_i$; and where $N(0, C_i^2\tau_i^2)$ covers $N(0, \tau_i^2)$ corresponds to $|\beta_i| > \delta_i$. From (2.8), if $\gamma_i = 1$, then $P(\beta_i|\gamma_i) = N(0, C_i^2\tau_i^2)$ and if $\gamma_i = 0$, then $P(\beta_i|\gamma_i) = N(0, \tau_i^2)$. So this suggests that if $|\beta_i| \leq \delta_i$, then $\beta_i = 0$ and X_i may be excluded from the model. Since the SSVS approach selects variables based on practical significance not on statistical significance, the largest value δ_i of β_i for setting $\beta_i = 0$ makes no practical difference. An easy way to select δ_i is to use the ratio of change in Y and X_i. Assume ΔY is the amount changed (which is insignificant) and ΔX_i is the amount changed (which is significant), then let $\delta_i = \Delta Y / \Delta X_i$. Any β_i smaller than δ_i would be too insignificant to be included in the model and this choice does not depend on γ.

After the choice of δ_i, τ_i and C_i are selected. The choice of τ_i and C_i can make $P(\beta_i|\gamma_i = 0) = N(0, \tau_i^2)$ cover $P(\beta_i|\gamma_i = 1) = N(0, C_i^2 \tau_i^2)$ exactly on the interval $(-\delta_i . \delta_i)$. And this can hold if τ_i^2 and C_i^2 satisfy

$$\frac{\ln(C_i^2 \tau_i^2 / \tau_i^2)}{1/\tau_i^2 - 1/C_i^2 \tau_i^2} = \delta_i^2 \tag{2.17}$$

Or it can be written as

$$\frac{\ln C_i^2}{1/\tau_i^2 (1 - 1/C_i^2)} = \delta_i^2$$

so

$$\tau_1 = \left[\frac{C_i^2 - 1}{2\ln(C_i) \bullet C_i^2} \right]^{1/2} \bullet \delta_i \tag{2.18}$$

In theory, C_i can be chosen arbitrarily . However if the value of C_i is too large, computational problems arise. George and McCulloch (1993,1995 & 1997) suggest choosing a C_i which is less than 10,000.

The final step for setting up the hierarchical model is to choose R_γ, the prior correlation matrix, conditional on γ. If we assume that the components of β are independent under $P(\beta|\gamma)$ then R_γ could be set as I_n and this is a very simple choice (George et al. 1995).

The other choice is to use the correlation structure based on the regressor cross product matrix $(X_i'X_i)^{-1}$ (George et al. 1993,1995 &1997). Now the hierarchical set-up is complete; the next step is to use Gibbs sampling to perform the SSVS.

2.2.3 SSVS by Gibbs Sampling

From (**2.2.2**), the goal is to obtain the posterior distribution $P(\gamma|Y)$ which contains useful information for variable selection; but $P(\gamma|Y)$ may not be analytically evaluated because of the difficulty of integration. The SSVS does not require computation of the entire 2^k possible posterior probabilities in $P(\gamma|Y)$, but rather, the algorithm searches promising subsets of the model space. This is supposed to improve computational speed..

SSVS uses the Gibbs sampling method to generate a sequence

$$\gamma^1, \gamma^2, \cdots\cdots\cdots \tag{2.19}$$

which will converge to γ ; $P(\gamma|Y)$. Those γ with high posterior probability will appear frequently in the samples and thus are easy to identify. In fact, most of the 2^k elements of the γ are very small indicating that they have a very small probability and thus can be discarded.

SSVS generates the sequence (2.19) to the full conditional distribution $P(\beta, \sigma, \gamma|Y)$ and this will generate a complete sequence of parameters

$$\beta^{(0)}, \sigma^{(0)}, \gamma^{(0)}, \beta^{(1)}, \sigma^{(1)}, \gamma^{(1)}, \cdots\cdots\cdots\cdots \tag{2.20}$$

This sequence converges to $P(\beta, \sigma, \gamma | Y)$, according to the Markov Chain theorem. The initial choice of $\beta^{(0)}$ and $\gamma^{(0)}$ could be based on the results from least squares or stepwise regressions. A conservative starting value for $\gamma^{(0)}$ could be to set $\gamma = (1,1,1,\cdots\cdots,1)$ which reflects the belief that all variables be included in the model. Another choice is to set $\gamma^{(0)} = (0,0,\cdots\cdots,0)$ which is likely to lead to much more parsimonious models. Initial values of the parameters, $\beta^{(0)}, \sigma^{(0)}, \gamma^{(0)}$, are chosen and subsequent values $(\beta^{(1)}, \sigma^{(1)}, \gamma^{(1)})$ are obtained using the sequence:

$$\beta^{(0)} \text{ into } P(\beta | \sigma^2, \gamma, Y)$$

$$\sigma^{(0)} \text{ into } P(\sigma^2 | \beta, \gamma, Y)$$

$$\gamma^{(0)} \text{ into } P(\gamma_i | \beta, \sigma^2, \gamma_1, \gamma_2, \cdots\cdots, \gamma_{i-1}, \gamma_{i+1}, \cdots\cdots, \gamma_k, Y) \tag{2.21}$$

Because of the hierarchical structure of prior distribution, the conditional distribution of σ only depends on β and Y, and the conditional probability of γ_i only depends on β and $\gamma_{-i} = (\gamma_1, \gamma_2, ------, \gamma_{i-1}, \gamma_{i+1}, ---- \gamma_k)$. So (2.21) can be obtained as

$$\beta^{(0)} \text{ into } P(\beta | \sigma^2, \gamma, Y)$$

$$\sigma^{(0)} \text{ into } P(\sigma^2 | \beta, \gamma, Y) = P(\sigma^2 | \beta, Y)$$

$$\gamma^{(0)} \text{ into } P(\gamma_i | \beta, \sigma^2, \gamma_{-i}, Y) = P(\gamma_i | \beta, \gamma_{-i}) \tag{2.22}$$

Subsequent values of β, σ^2, γ are obtained by iterating, substituting the most recent values from the sequence into (2.22).

Next samples are drawn from these three conditional distributions. The SSVS algorithm will be efficient and converge quickly if these conditionals have standard distributions. The first step is to draw β successively from

$$P(\beta|\sigma^2,\gamma,Y) = N_k((X'X+\sigma^2(D_\gamma R_\gamma D_\gamma)^{-1})^{-1}),\sigma^2(X'X+(D_\gamma R_\gamma D_\gamma)^{-1})^{-1}) \qquad (2.23)$$

This step requires an update of $(X'X+\sigma^2(D_\gamma R_\gamma D_\gamma)^{-1})^{-1}$ every time a new σ^2 and γ are used. Updating can be easily done in the following way using a Cholesky decomposition (Thisted, 1988 and Gelman et al. 1998). From Thisted (1988, p118), let $W=\sigma D_r^{-1}$ which is an $1\times p$ vector and $(X'X+\sigma^2 D_r^{-2})$ (assume $R_r = I_n$, which is independent of γ) can be decomposed as

$$\begin{pmatrix} X \\ \sigma D_r^{-1} \end{pmatrix}' \begin{pmatrix} X \\ \sigma D_r^{-1} \end{pmatrix} \qquad (2.23\text{-}1)$$

So when a new σ and D_r^{-1} are generated, we substitute the new values into (2.23-1) and update the cross product term in (2.23). The case with R_r P $(X'X)^{-1}$ is similar but requires a QR decomposition for X; updating the algorithm is then the same as (2.23-1).

Second, draw σ^2 from

$$P(\sigma^2|\beta,\gamma,Y) = P(\sigma^2|\beta,Y) = InvG(\frac{n+\lambda_\gamma}{2},\frac{|Y-X\beta|^2+\upsilon\lambda_\gamma}{2}) \qquad (2.24)$$

the updated inverse Gamma distribution of (2.14).

Finally, draw γ by sampling each γ_i successively from the Bernoulli distribution with probability $\frac{a}{a+b}$ where

$$P(\gamma_i = 1|\beta,\gamma_{-i}) = \frac{a}{a+b}, \qquad (2.25)$$

$a = P(\beta|\gamma_{-i},\gamma_i = 1) \bullet P(\gamma_{-i},\gamma_i = 1)$, and $b = P(\beta|\gamma_{-i},\gamma_i = 0) \bullet P(\gamma_{-i},\gamma_i = 0)$.

Using the independence assumption in the Bernoulli prior (2.15), (2.25) can be rewritten as

$$P(\gamma_i = 1|\beta_i) = \frac{a}{a+b} \tag{2.26}$$

where $a = P(\beta_i|\gamma_i = 1) \bullet \theta_i$ and $b = P(\beta_i|\gamma_i = 0) \bullet (1-\theta_i)$

As can be seen from (2.26), a is the product of (2.8) and (2.9) when $\gamma_i = 1$, while b is the product of (2.8) and (2.9) when $\gamma_i = 0$. Also, each time a new value of γ is generated, the value of γ_i (not all, but some) is decided randomly from (2.25). Since γ_i is the indicator for X_i being included or excluded from the model, generation of the sequence (2.20) is equivalent to performing the stochastic search procedure. This is the main objective of SSVS, which is to search for high frequencies (i.e. which model appears most) rather than to evaluate the entire posterior probability.

2.3 The Alternative Approach (MBVS)

It was mentioned in the first chapter that there is a similar approach to Bayesian model selection called MBVS (Brown et al. (1998). As a multivariate procedure, the MBVS is generalized to consider p regressors and q responses (dependent variables). Like SSVS, MBVS uses a latent vector to identify two types of regression coefficients: those close to 0 and those not. Brown et al. derive the marginal distribution for this binary latent vector and use the MCMC approach (like Gibbs sampling) to draw samples form the posterior distribution of the known parameters. This alternative approach can be used to select appropriate variables from a large number of regressors by using MCMC, and to approximate the posterior distribution of the binary latent vector directly.

In the current context the MVBS can be simplified to a univariate case by letting $q = 1$ with response $Y = (Y_1)$ which depends conditionally on p independent variables $x = (x_1, ---, x_p)'$. For the responses, $l = 1, Y_l$ is assumed to have a mean that is $\eta(\alpha_l + \beta_l' x)$, where $\eta(\cdot)$ is a known continous fuction. The β_l is a $p-$ vector of unknown slope parameters and α_l is an unknown scale parameter. For n independent observations $Y_l(q \times 1)$, conditional on $x_i(p \times 1), i = 1,2 --, n$, this is a multivariate generalized linear model. The intercept vector, α, is ($q \times 1$), the slope matrix $B = (\beta_1, ---, \beta_q)(p \times q)$, with covariance matrix, Σ. Then the joint prior distribution for the α, B and Σ can be decomposed as

$$\pi(\alpha, B, \Sigma) = \pi(\alpha|\Sigma)\pi(B|\Sigma)\pi(\Sigma)$$

which assumes independence between α and B. The latent binary vector is denoted as γ where the j th element of γ could be 0 or 1 (similar to SSVS) and this vector is associated with $\pi(B|\Sigma)$. Like SSVS, when the j th element is equal to 0 then the j th independent variable can be deleted from the model. So $\pi(B|\Sigma)$ can be elaborated as

$$\pi(B, \gamma|\Sigma) = \pi(B|\Sigma, \gamma)\pi(\gamma) \tag{2.3.1}$$

The main goal for this approach is to evaluate the posterior distribution for the latent vector γ conditional on X and Y, that is, $\pi(\gamma|X, Y)$. And this could be approximated as

$$\pi(\gamma|X, Y) \propto \pi(\gamma)\iiint f(Y|X, \alpha, B, \Sigma)\pi(\alpha|\Sigma)\pi(B|\Sigma, \gamma)\pi(\Sigma)d\alpha d\beta d\Sigma \tag{2.3.2}$$

where $f(Y|X, \alpha, B, \Sigma)$ is the likelihood function. Next is the model setting for the MBVS.

2.3.1 Model Settings for MBVS

Assume conditional on $\alpha, B, \gamma, \Sigma$, the standard multivariate regression model is

$$Y - 1\alpha' - XB \,;\, N(I_n, \Sigma) \tag{2.3.3}$$

with $n \times q$ random matrix Y, 1 is an $n \times 1$ vector of 1s, X is an $n \times p$ and B is $p \times q$ matrix of regression coefficients. The prior distribution for α, given Σ, is

$$\alpha - \alpha_0' \,;\, N(h, \Sigma) \tag{2.3.4}$$

And given Σ and γ, the prior distribution for B, is

$$B - B_0 \,;\, N(H_\gamma, \Sigma) \tag{2.3.5}$$

The prior for Σ is assumed distributed as inverse Wishart distribution which is

$$\Sigma \,;\, IW(\delta, Q) \tag{2.3.6}$$

where δ and Q are scale parameters.

The prior for γ is assumed to be the Bernoulli distribution, that is,

$$P(\gamma_j = 1) = w_j \text{ and } P(\gamma_j = 0) = 1 - w_j \tag{2.3.7}$$

with w_j to be specified by the researcher.

For the H_γ, one could use the assumption of George et al. (1993)

$$H_\gamma = D_\gamma R_\gamma D_\gamma \tag{2.3.8}$$

where D_γ is a diagnoal matrix and R_γ is a correalation matrix. The *jth* element of D_γ^2 is taken to be v_{0_j} where $\gamma_j = 0$ and v_{1_j} where $\gamma_j = 1$. A special case is often considered where $R_r = I$, the identity matrix, with $B_0 = 0$, a zero matrix. With this prior setting, $\gamma_j = 0$ indicates that the *jth* row of B has zero variance and when $\gamma_j = 1$ indicates that the

jth row has non-zero variance. For this setting, the prior distribution of B is a matrix where each column has a singular p_r dimensional distribution, so the prior distribution for B becomes

$$B_{(\gamma)} - B_{0(\gamma)} ; N(H_{(\gamma)}, \Sigma) \tag{2.3.9}$$

Hence, $B_{(\gamma)}$ selects row of B that have $\gamma_j = 1$.

The hyperparameter h of the prior distribution of α can be set to a large value which makes α_0 irrelevant. The hyperparameter Q of the prior distribution of Σ , is given a simple form, kI_q, after scaling of the dependent variables. For weak prior information, set $\delta = 3$ when $E(\Sigma) = Q/(\delta - 2) = Q$. So $\delta = 3$ is the smallest value for the expectation of Σ exists which is also convenient for us. The next step is to derive the posterior distribution for $\pi(\gamma | X, Y)$.

2.3.2 The Posterior Distribution

The p.d.f. of Y is

$$f_Y(Y|\alpha, B, \Sigma) = c(n,q)|\Sigma|\exp[-\frac{1}{2}tr((Y - 1\alpha' - XB)\Sigma^{-1}(Y - 1\alpha' - XB)'))] \qquad 2.3.10)$$

where $c(n,q)$ is a constant. Now assume the columns of Y and X have been centered by subtracting their columns mean. That is

$$\bar{Y}_l = 0, \; l = 1,2,---,q$$

$$\bar{x}_j = 0, j = 1,2,---,p \tag{2.3.11}$$

The joint p.d.f. of (α, B) given Σ, γ, is

$$\pi(\alpha|\Sigma) \text{ P } h^{-q/2}|\Sigma|\exp[-\frac{1}{2h}(\alpha - \alpha_0)'\Sigma^{-1}(\alpha - \alpha_0)] \tag{2.3.12}$$

and

$$\pi(B|\Sigma,\gamma)\,\mathrm{P}\ |H_\gamma|^{-q/2}|\Sigma|^{-p/2}\exp[-\frac{1}{2}tr(H_\gamma^{-1}(B-B_0)\Sigma^{-1}(B-B_0)')] \tag{2.3.13}$$

After integrating over (α, B) given Σ and γ, and integrating B given Σ and γ, the posterior distribution of γ is approximated as

$$\pi(\gamma|X,Y)\,\mathrm{P}\ g(\gamma) = (|H_\gamma||K_\gamma|)^{-q/2}|Q_\gamma|^{-(n+\delta+q-1)/2}\,\pi(\gamma) \tag{2.3.14}$$

where

$$\begin{aligned} Q_\gamma &= Q + C - M'K_\gamma^{-1}M \\ &= Q + Y'Y - Y'XK_\gamma^{-1}X'Y \end{aligned} \tag{2.3.15}$$

and

$$M = X'Y + H_\gamma^{-1}B_0, \tag{2.3.16}$$

$$C = Y'Y + B_0'H_\gamma^{-1}B_0, \tag{2.3.17}$$

$$K_\gamma = X'X + H_\gamma^{-1} \tag{2.3.18}$$

for $B_0 = 0$. The computation for the posterior distribution can be derived directly from the equation once the hyperparameters (H_γ, Q and δ) have been specified.

2.3.3 Prior Settings and Updating

The prior distribution for this approach requires B, Σ, γ and the Bernoulli distribution $\pi(\gamma)$. The prior distribution for B given γ depends on H_γ. Letting $R_\gamma = I$ will simplify the process. One alternative automatic prior when $v_{0j} = 0$ is

$$H_{(\gamma)} = c(X_\gamma' X_{(\gamma)}) \tag{2.3.19}$$

which implies that the subset $X_{(\gamma)}$ of columns of X chosen to correspond to $\gamma_j = 1$ is of full rank.

Now consider both cases when $v_{0j} > 0$ as well as $v_{0j} = 0$. The first part of the equation (2.3.14) can be written as

$$\begin{aligned} |H_\gamma||K_\gamma| &= |H_\gamma||X'X + H_\gamma^{-1}| \\ &= |H_\gamma^{1/2}X'XH_\gamma^{1/2} + I| \\ &= |\tilde{X}'\tilde{X}| \end{aligned} \tag{2.3.20}$$

where

$$\tilde{X} = \begin{pmatrix} XH_\gamma^{1/2} \\ I_p \end{pmatrix}$$

is an $(n+p) \times p$ matrix. And,

$$\tilde{Y} = \begin{pmatrix} Y \\ 0 \end{pmatrix}$$

is an $(n+p) \times q$ matrix. So Q_γ from equation becomes Q plus

$$\tilde{Y}'\tilde{Y} - \tilde{Y}'\tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{Y} \tag{2.3.21}$$

and this is the residual sum of square matrix from the least squares of $\tilde{Y}$ on $\tilde{X}$. The computation will simplified further if $\tilde{X}$ is reduced to the $(n+p) \times p_\gamma$ matrix $\tilde{X}_{(\gamma)}$, which is the special form of $\tilde{X}$ (when selecting the $\gamma_j = 1$ columns). The QR – decomposition of $(\tilde{X}_{(\gamma)}, \tilde{Y})$ is given examples in Seber (1984) with updating algorithms *qrinsert* and *qrdelete* used to add or delete columns to the reduced

$(n+p) \times p_\gamma$ matrix. After these simplification with automatic prior of $H_{(\gamma)}$,the relevant part of Q_γ becomes

$$Y'Y - Y'X_{(\gamma)}\{X'_{(\gamma)}X_{(\gamma)} + (1/c)X'_{(\gamma)}X_{(\gamma)}\}^{-1}X'_{(\gamma)}Y$$

$$= \{c/(c+1)\}\{Y'Y - Y'X_{(\gamma)}(X'_{(\gamma)}X_{(\gamma)})^{-1}X'_{(\gamma)}Y\} + Y'Y/(c+1) \qquad (2.3.22)$$

and the required quantities for the equation can be obtained by regressing Y on $X_{(\gamma)}$. For $|H_{(\gamma)}K_{(\gamma)}|$ of the equation, will be simplified to $(c+1)^{p_\gamma}$. This fast algorithm allows one to update the quantities of the posterior distribution more efficiently. Next, the MCMC method used to carry out the computations is discussed.

2.3.4. Computations Using MCMC

Although one can analyze the posterior distribution directly by equation, the right hand side of the equation would require the computation of all possible 2^p subsets of the vector γ. Using current microcomputing power, this becomes infeasible when the independent variables are more than 20. By using the MCMC sampling method, the higher marginal probabilities of γ_j can be identified and promising models will be selected.

The simple Gibbs sampling algorithm will be implemented by generating random samples from the conditional distribution

$$\gamma_j | \gamma_{-j}, Y, X \quad j = 1,2 - —, p, \qquad (2.3.23)$$

where $\gamma_{-j} = (\gamma_1, \gamma_2, - - -, \gamma_{j-1}, \gamma_{j+1}, - - \gamma_p)$. Then, as in SSVS, the random samples will be drawn from the conditional Bernoulli distribution with probability

$\theta_j / (\theta_j + 1)$ where

$$\theta_j = g(\gamma_j = 1, \gamma_{-j} | Y, X) / g(\gamma_j = 0, \gamma_{-j} | Y, X). \qquad (2.3.24)$$

MBVS avoids the computation of the conditional posterior distribution for every parameter. It becomes an appropriate approach when the computational burden of SSVS is too great. For this reason, MVBS is used chapter four in lieu of SSVS in order to ascertain how it compares to traditional model selection procedures.

2.4 Geweke's Approach

Geweke (1994) proposes a subjective prior approach to solve the variable selection problem. As in the usual regression set-up,

$$Y = X\beta + \varepsilon \qquad \varepsilon \text{ ; } N(0, \sigma^2 I_n) \qquad (2.27)$$

where Y is an $n \times 1$ vector containing the dependent variables and X is an $n \times k$ matrix of independent variables. Geweke assumes that k^* out of k parameters have nonzero coefficients with prior probability 1, and there is a positive probability that the remaining $k - k^*$ variables have coefficients equal to zero. Along with (2.27), Geweke also assumes the prior distributions (for parameters) are all mutually independent, but this assumption may be weakened under special conditions. The other assumption is that the prior distribution of each coefficient is a mixture of normal or truncated normal distribution and discrete mass at 0. Under these assumptions, the prior distribution can be formulated as follows.

For prior probability $P_i, \beta_i = 0$, but conditional on $\beta_i \neq 0$ the prior distribution for β_i is a truncated normal distribution in interval $(\lambda_i, \mathrm{v}_i)$ which is $TN_{(\lambda_i, \mathrm{v}_i)}(\underline{\beta_i}, \tau_i^2)$ and the joint prior distribution for β_i can be written as:

$$P(\beta_1,\beta_2,\ldots,\beta_k)=\prod_{i=1}^{k}P(\beta_i)=P_i h_i(\beta_i)+(1-P_i)(2\pi)^{\frac{-1}{2}}\tau_i^2\left[\Phi\left(\frac{v_i-\underline{\underline{\beta}}_i}{\tau_i}\right)-\Phi\left(\frac{\lambda_i-\underline{\underline{\beta}}_i}{\tau_i}\right)\right]^{-1}$$

$$EXP\left[\frac{(\beta_i-\underline{\underline{\beta}}_i)}{2\tau_i^2}\right]\bullet I_{(\lambda_i,v_i)}(\beta_i) \tag{2.28}$$

where $P(\beta_i)$ is the prior p.d.f. of β_i; h(x) = 0 if x < 0 and h(x) = 1 if x > 0; $I_s(x)=1$ if $x\in s$ and $I_s(x)=0$ if $x\notin s$. $\Phi(\ )$ is the c.d.f. of the standard normal distribution; $0<\tau_i<\infty$; $-\infty<\lambda_i\le v_i<\infty$. The prior distribution for σ is

$$\underline{v}\,\underline{\sigma^2};\ \chi^2(\underline{v}) \tag{2.29}$$

So from (2.28), we can see that this is a combination of two distributions, one of which has mass at 0 and the other has a truncated normal distribution. Geweke argues that with this prior set-up, it is easy to eliminate the objective prior about the coefficient and it is also easy to compute.

Again, the computational procedure implemented here is Gibbs sampling with complete blocking. It proceeds as follows. Draw each β_i, $i=1,\cdots,k$, from its conditional distribution (conditional on $\beta_l\,(l\neq i)$ and σ), and then draw σ from its conditional distribution (conditional on β). The algorithm to obtain the conditional distribution is as follows:

1. Use the ordinary least square to obtain $\underline{\beta}_l=(X'X)^{-1}X'Y$, then get the residual

$$Z_i=Y_i-\sum\nolimits_{l\neq j}\beta_l X_{il} \tag{2.30}$$

the conditional distribution of β_j follows from

$$Z_i=\beta_j X_{ij}+\varepsilon_i \quad \text{where } \varepsilon_i;N(0,\sigma^2)\ i=l,k,n$$

2. Compute the estimate of the omitted coefficient

$$b = \sum_{i=1}^{n} X_{ij} Z_i / \sum_{i=1}^{n} X_{ij}^2 \tag{2.31}$$

and the precision is

$$W^2 = \sigma^2 / \sum_{i=1}^{n} X_{ij}^2 \tag{2.32}$$

(2.30) and (2.31) are just the usual form of least square for β and σ.

3. The kernel for likelihood function $l(\beta, \sigma)$ is

$$EXP\left[-\sum_{i=1}^{n} \left(Z_i - \beta_j X_{ij} \right)^2 / 2\alpha^2 \right] = EXP\left[-\sum_{i=1}^{n} Z_i^2 / 2\alpha^2 \right] \tag{2.33}$$

conditional on $\beta_j \neq 0$ and $\beta_{\mathrm{j}} = 0$, (2.33) will become

$$EXP\left[-\sum_{i=1}^{n} Z_i^2 / 2\alpha^2 \right] \text{ condition on } \beta_j = 0 \tag{2.34}$$

and

$$\begin{aligned} &EXP\left[-\sum_{i=1}^{n} \left(Z_i - \underline{\beta}_j \right)^2 / 2\alpha^2 \right] \bullet \\ &(2\pi)^{-1/2} \tau_j^{-1} \left[\Phi\left(v_j - \underline{\beta}_j \right) / \tau_j - \Phi\left(\lambda_j - \underline{\beta}_j \right) / \tau_j \right]^{-1} \bullet \\ &EXP\left[-\left(\beta_j - \underline{\beta}_j \right)^2 / 2\tau_j^2 \right] I_{(\lambda_j, v_j)}\left(\beta_j \right) \text{condition on } \beta_j \neq 0 \end{aligned} \tag{2.35}$$

insert (2.31) and (2.32) into (2.35); it becomes

$$\begin{aligned} &EXP\left[-\sum_{i=1}^{n} \left(Z_i - bX_{ij} \right)^2 / 2\alpha^2 \right] EXP\left[-\left(\beta_j - b \right)^2 / 2W^2 - \left(\beta_j - \underline{\beta}_j \right)^2 / 2\tau_j^2 \right] \bullet \\ &(2\pi)^{-1/2} \tau_j^{-1} \left[\left[\Phi\left(v_j - \underline{\beta}_j \right) / \tau_j \right] - \left[\Phi\left(\lambda_j - \underline{\beta}_j \right) / \tau_j \right] \right]^{-1} I_{(\lambda_j, v_j)}\left(\beta_j \right) \end{aligned} \tag{2.36}$$

4. Compute the weight on W and $\underline{\beta}_j$

$$\sigma_1^2 = (W^{-2} + \tau_j^{-2})^{-1} \tag{2.37}$$

$$\overline{\beta_j} = \sigma_1^2 (W^{-2} b + \tau_j^{-2} \underline{\beta}_j)^{-1} \tag{2.38}$$

(2.37) is the pool estimate weighted by W and τ_j, where (2.38) is the weighted average for $\underline{\beta}_j$ and b (weighted by W^{-2} and τ_j^{-2}).

For truncated normal distribution, insert (2.37) and (2.38) into (2.36), and it will become

$$\begin{aligned}&EXP\left[-\sum_{i=1}^{n}\left(Z_i - bX_{ij}\right)^2/2\alpha^2\right]EXP\left[-\left(\beta_j-\underline{\beta}_j\right)^2/2\alpha^2\right]\bullet\\&EXP\left[-\left(b^2/2W^2+\beta_j^2-\overline{\beta}_j/2\alpha_1^2\right)\right]\bullet\left(2\pi\right)^{-1/2}\bullet\\&\tau_j^{-1}\left[\Phi\left(\frac{v_j-\underline{\beta}_j}{\tau_j}\right)-\Phi\frac{\lambda_j-\underline{\beta}_j}{\tau_j}\right]^{-1}I_{(\lambda_j,v_j)}\left(\beta_j\right)\end{aligned} \tag{2.39}$$

5. Integrate (2.39) over β_j to remove the conditional on $\beta_j = 0$ or $\beta_j \neq 0$. Then (2.39) will become

$$\begin{aligned}&EXP\left[-\sum_{i=1}^{n}\left(Z_i - bX_{ij}\right)^2/2\alpha^2\right]EXP\left[-\left(b/2W^2+\underline{\beta}_j^2/2\tau_j^2-\overline{\beta}_j^2/2\alpha_1^2\right)\right]\bullet\\&\left(\alpha_1/\tau_j\right)\bullet\left[\left[\Phi\left(\frac{v_j-\overline{\beta}_j}{\alpha_j}\right)-\Phi\left(\frac{\lambda_j-\overline{\beta}_j}{\alpha_j}\right)\right]\bullet\left[\Phi\left(\frac{\lambda_j-\underline{\beta}_j}{\tau_j}\right)-\Phi\left(\frac{v_j-\underline{\beta}_j}{\tau_j}\right)\right]\right]^{-1}\end{aligned} \tag{2.40}$$

6. Compute the conditional Bayes factor, favor of $\beta_j \neq 0$ against $\beta_j = 0$ by taking the ratio of equation (2.33) and (2.40).

Thus BF =

$$
\begin{aligned}
&EXP\left[\sum_{i=1}^{n} Z_i^2 - \sum_{i=1}^{n}\left(Z_i - bX_{ij}\right)^2 / 2\alpha^2\right] \bullet EXP\left[-\left(b/2W^2 + \underline{\beta}_j^2/2\tau_j^2 - \overline{\beta}_j^2/2\alpha_1^2\right)\right] \bullet \\
&\left(\alpha_1/\tau_j\right) \bullet \left[\left[\Phi\left(\frac{v_j - \overline{\beta}_j}{\alpha_j}\right) - \Phi\left(\frac{\lambda_j - \overline{\beta}_j}{\alpha_j}\right)\right] \bullet \left[\Phi\left(\frac{v_j - \underline{\beta}_j}{\tau_j}\right) - \Phi\left(\frac{\lambda_j - \underline{\beta}_j}{\tau_j}\right)\right]\right] \\
&= EXP\left[\overline{\beta}_j^2/2\alpha_1^2 - \underline{\beta}_j^2/2\tau_j^2\right] \bullet \left(\alpha_1/\tau_j\right) \bullet \\
&\left[\left[\Phi\left(\frac{v_j - \overline{\beta}_j}{\alpha_j}\right) - \Phi\left(\frac{\lambda_j - \overline{\beta}_j}{\alpha_j}\right)\right] \bullet \left[\Phi\left(\frac{v_j - \underline{\beta}_j}{\tau_j}\right) - \Phi\left(\frac{\lambda_j - \underline{\beta}_j}{\tau_j}\right)\right]\right]
\end{aligned}
\tag{2.41}
$$

7. Compute the conditional posterior probability for $\beta_j = 0$ by using Bayes factor

$$\overline{P}_j = \frac{\underline{P}_j}{\underline{P}_j + (1 - \underline{P}_j) \bullet BF} \tag{2.42}$$

(2.42) is the ratio of prior probability of $\beta_j = 0$ by using the weighted average of two prior probabilities ($\beta_j \neq 0$ and $\beta_j = 0$), where weighted by the Bayes factor. When the Bayes factor approaches one, $\overline{P}_j = \underline{P}_j$, the evidence strongly supports the hypothesis that $\beta_j \neq 0$. But as BF approaches zero, $\overline{P}_j = 1$, which supports the hypothesis that $\beta_j = 0$.

8. After computing (2.42), decide whether $\beta_j = 0$ or $\beta_j \neq 0$. The decision is made by drawing a random value, μ from a uniform distribution on [0,1]. If $\overline{P}_j < \mu$, then $\beta_j \neq 0$ and draw β_j from the truncated normal distribution $TN_{(\lambda_j, v_j)}(\overline{\beta}_j, \sigma_1^2)$. If $\overline{P}_j \geq \mu$, then pick $\beta_j = 0$ and exclude X_j from the model.

9. The posterior distribution under β_j for σ is

$$\left[\underline{\mathrm{v}\sigma}^2 + (Y - X\beta)'(Y - X\beta)\right] / \sigma^2 \; ; \chi^2_{(\nu+n)} \tag{2.43}$$

The above nine steps are Geweke's algorithms to compute the posterior probability. Gibbs sampling is carried out in the usual way. First, set the initial value for $\beta^{(0)} = (\beta_1^{(0)}, \beta_2^{(0)}, \cdots\cdots\cdots, \beta_k^{(0)})$ and σ. The starting value could be the least square estimator or from stepwise regression, or drawn from their prior distribution. Second, draw β from its respective conditional posterior distribution (from the $TN_{(\lambda_j, \nu_j)}(\overline{\beta}_j, \sigma_1^2)$ if it is truncated); draw σ^2 from (2.43). The objective for this is to determine the posterior probability for 2^{k-k^*} models. The conditional posterior probability of $\beta_j = 0$, (2.42) can be an indicator for this Gibbs sampling when it proceeds. After each iteration, record $\overline{P}_j$ regarding $\beta_j = 0$ or record $(1 - \overline{P}_j)$ regarding $\beta_j \neq 0$. And the posterior probability for $\beta_j = 0$ can be the proportion of the Gibbs samples for which the coefficient is set equal to zero (that is, $\overline{P}_j > \mu$ in step 8).

But as Geweke (1994) argues, the degree of collinearity among independent variables has a serious effect on the convergence of Gibbs samplers. So assessing the severity of collinearity among independent variables is recommended. The higher the degree of collinearity, the more iterations needed to assure convergence of the Gibbs sampler.

2.5 The Bayesian Model Averaging Approach (BMA)

The last Bayesian approach of model selection considered is Bayesian Model Averaging (BMA). Like all Bayesian procedures, BMA combines information prior information with that from the data and the model to form the posterior probability of the model. It can be expressed as:

$$P(I|D) = \sum_{k=1}^{K} P(I| M_k, D) P(M_k | D) \tag{2.44}$$

where I is the quantity of interest, D is the data set and M_k is the kth model ($k = 1,2,3,\cdots\cdots, K$). The posterior probability for model M_k is given by (Raftery, Madigan and Hoeting, 1998)

$$P(M_k | D) = \frac{P(D| M_k) P(M_k)}{\sum_{l=1}^{k} P(D| M_l) P(M_l)} \tag{2.45}$$

where

$$P(D|M_k) = \int P(D|\theta_k, M_k) P(\theta_k|M_k) d\theta_k \tag{2.46}$$

is the marginal likelihood function of M_k, θ_k is the vector of parameters of the model, $P(\theta_k | M_k)$ is the prior probability for θ_k under model M_k, $P(D|\theta_k, M_k)$ is the likelihood function, and $P(M_k)$ is the prior probability for model M_k.

From (2.44), we can see BMA averages over the interest, I, under data and models using the conditional model probabilities as weights. It utilizes the information in a special way that accounts for the uncertainties associated with the models. The method to carry out BMA is adopted from Hoeting (1994), Raftery et al.(1997) and Hoeting et al.(1998).

The critical points to evaluate (2.44) are:

(1) Possible large number of terms in (2.44)

(2) Prior selection of $P(M_k)$, the model prior probability

(3) Integrals involved in (2.46)

The methods considered to solve (1) and (3) will be Occam's window (Madigan & Raftery, 1994) and MC^3 composition (Madigan & York, 1995). Next we set up the framework of BMA.

2.51 Framework for BMA

The model considered here is the ordinary linear regression model (2.4) which has form

$$Y = X\beta + \varepsilon$$

where Y, X, β and ε have been defined previously in (2.4).

The differences from the two previous approaches are that, BMA argues that the prior distribution needs to reflect the uncertainties about the parameters and assumes a reasonable prior constraint. By using the standard Normal-Gamma conjugate prior, β and σ^2 are assuming as:

$$\beta \; ; \; N(\mu, \sigma^2 V) \tag{2.47}$$

$$v\lambda \; ; \; \chi^2_v$$

$V, \lambda,$ $(K+1)\times(K+1)$ matrix V and $(K+1)\times 1$ vector μ needs to be chosen.

Assume each β_i in β is independent of the others and center the distribution of β around 0. Set $\mu = (\beta_0', 0, 0, \cdots\cdots, 0)$ where β_0' is the sample mean of dependent variable. The covariance matrix of β, V(β) is the covariance matrix which is

$$
V(\beta)=\begin{bmatrix}
S_Y^2 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \Phi^2 S_1^{-2} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & & \\
- & - & - & - & \Phi^2 S_i^{-2} & - & - & - & - & - \\
- & - & - & - & - & \Phi^2\left(\frac{1}{n}X'_i X'\right) & - & - & - & - \\
- & - & - & - & - & - & \Phi^2 S_{i+1}^{-2} & - & - & - \\
- & - & - & - & - & - & - & - & - & - \\
- & - & - & - & - & - & - & - & - & - \\
- & - & - & - & - & - & - & - & - & \Phi^2 S_K^{-2}
\end{bmatrix} \qquad (2.48)
$$

Where S_Y^2 is the sample variance of Y, S_i^2 is the sample variance of $X_i, i = 1,2,3,\cdots\cdots, K$, and Φ needs to be chosen. This prior covariance matrix for β is chosen to reflect the increasing precision for each β_i as the variance of X_i increases and will not be affected by the change of Y and X.

If the model has dummy variables (in this research, I will have several regional dummy variables), the prior variance of $\beta = (\beta_{i1}, \beta_{i2}, \cdots\cdots, \beta_{id})$ is set as $\sigma^2\Phi^2(\frac{1}{n}X_i' X_i)^{-1}$, where X_i is the n x d matrix for dummy variables and each dummy variable is centered around its sample mean. This prior set-up is related to the special prior called g prior from Zellner (1986).

Next select parameters λ, v and Φ, for which some criteria need to be defined. Assuming all variables are standardized as N (0,1), we define the following criteria:

1) $P(\sigma^2 \le 1)$ is the large, that means consistent variance is less than 1
2) $P(\sigma^2)$, prior probability for σ^2, is flat over $(\alpha,1)$ for reasonable small α. This conflicts with 1), but assures that $P(\sigma^2)$ distributed over $(\alpha,1)$.

3) $P(\beta_1, \beta_2, \cdots\cdots, \beta_k)$, prior probability of $P(\beta)$, is reasonably flat over unit hypercube $[-1,1]^k$.

These three criteria are defined so the selection of λ, v and Φ can be consistent with the prior settings. Then, maximize $P(\sigma^2 \leq 1)$ subject to (Hoeting, 1994)

I) $$\frac{P(\beta_1 = 0, \beta_2 = 0, \cdots\cdots, \beta_k = 0)}{P(\beta_1 = 1, \beta_2 = 1, \cdots\cdots, \beta_k = 1)} \leq M_1 \qquad (2.49)$$

since

$$\begin{aligned} P(\beta) &= \int \frac{P(\beta, \alpha^2)}{P(\alpha^2)} \bullet P(\alpha^2) d(\alpha^2) \\ &= \int P(\beta | \alpha^2) \bullet P(\alpha^2) d(\alpha^2) \end{aligned} \qquad (2.50)$$

Again, following Hoeting(1994), (2.47) and (2.48) implies:

$$\begin{aligned} P(\alpha^2) &\propto (\alpha^2)^{-(v/2+1)} EXP[-v\lambda / 2\alpha^2] \\ P(\beta | \alpha^2) &\propto (\alpha^2)^{-k/2} EXP\left[\frac{-1}{2\alpha^2 \Phi^2} \beta'\beta\right] \end{aligned} \qquad (2.51)$$

Inserting (2.51) into (2.50) yields

$$\begin{aligned} P(\beta) &= \int (\alpha^2)^{-k/2} EXP\left[\frac{-1}{2\alpha^2\Phi^2}\beta'\beta\right] \bullet (\alpha^2)^{\left(\frac{-v}{2}+1\right)} EXP\left[\frac{-v\lambda}{2\alpha^2}\right] d(\alpha^2) \\ &= \int (\alpha^2)^{-\left(\frac{k+v}{2}+1\right)} EXP\left[\frac{-1}{2\alpha^2}\left[\frac{\beta'\beta}{\Phi^2} + v\lambda\right]\right] d(\alpha^2) \\ &\propto \left(\frac{1}{2\Phi^2}\sum_{i=1}^{k}\beta_i^2 + v\lambda / 2\right)^{-\left(\frac{v+\lambda}{2}\right)} \end{aligned} \qquad (2.52)$$

So (2.49) can be formed as

$$\frac{P(\beta_1 = 0, \beta_2 = 0, -----, \beta_k = 0)}{P(\beta \ = 1, \beta_2 = 1, -----, \beta_k = 1)} = \frac{P(\beta = 0)}{P(\beta = 1)}$$

$$=\left(\frac{\frac{1}{2\Phi^2}\bullet 0+\frac{\lambda v}{2}}{\frac{1}{2\alpha^2}\bullet k+\frac{\lambda v}{2}}\right)^{-\left(\frac{v+k}{2}\right)}$$

$$=\left(1+\frac{k}{\Phi^2 \lambda v}\right)^{\frac{v+k}{2}}=\overline{M}_1 \tag{2.53}$$

II) Let $P(\sigma^2)$ be reasonably flat over $(\alpha,1)$ for some small α

so

$$\frac{Max_{\alpha\le\sigma^2\le 1} P(\sigma^2)}{P(\sigma^2=\alpha)}\le \overline{M}_2 \tag{2.54}$$

and

$$\frac{Max_{\alpha\le\sigma^2\le 1} P(\sigma^2)}{P(\sigma^2)}\le \overline{M}_2 \tag{2.55}$$

Let S = σ^2 where $r=1/S$ is the precision and $r;\ \chi^2_v/\lambda_v$ from (2.47), so

$$P(S)=S^{-(v/2+1)}\bullet EXP[-v\lambda/2S]$$

this will be maximized when S = MLE, where MLE = $v\lambda/(2+v)$.

So (2.54) and (2.55) becomes

$$\frac{Max_{\alpha\le\sigma^2\le 1} P(\sigma^2)}{P(\sigma^2=\alpha)}=\frac{\left(\frac{v\lambda}{2+v}\right)^{-\left(\frac{v}{2}+1\right)}\bullet EXP\left[\frac{-v\lambda}{2\frac{v\lambda}{2+v}}\right]}{\alpha^{-\left(\frac{v}{2}+1\right)}\bullet EXP\left[\frac{-v\lambda}{2\alpha}\right]}$$

$$=\left(\frac{v\lambda}{2+v}\right)^{-\left(\frac{v}{2}+1\right)}\bullet EXP\left[-\left(\frac{v}{2}+1\right)\bullet\left(\frac{v\lambda}{2\alpha}\right)\right] \tag{2.56}$$

$$=\left(\frac{v\lambda e}{2+v}\right)^{-\left(\frac{v}{2}+1\right)}\bullet EXP\left[\frac{-v\lambda}{2\alpha}\right]$$

Also

$$\frac{Max_{\alpha \leq \sigma^2 \leq 1} P(\sigma^2)}{P(\sigma^2 = 1)} = \left(\frac{v\lambda e}{2+v}\right)^{-\left(\frac{v}{2}+1\right)} \bullet EXP\left[\frac{-v\lambda}{2}\right] \tag{2.57}$$

The reasonable range for v and λ suggested by Hoeting (1994) is

$$2\left(\frac{\alpha}{2-\alpha}\right)^2 \leq v \leq \infty$$

and

$$\alpha \leq \lambda \leq 2$$

This range is consistent with the assumption that $P(\sigma^2)$ is flat over $[\alpha, 1]$ and variables are standardized to be N (0,1).

So under different combinations of $\alpha, \overline{M}_1, \overline{M}_2$, and subject to (2.53), (2.54) and (2.55), $P(\sigma^2 \leq 1)$ can be maximized because different k will obtain different Φ and $MaxP(\sigma^2 \leq 1)$.

2.5.2 Computation of BMA

Here we consider two methods to carry out BMA. The first is the Occam's window algorithm from Madigan and Raftery (1994). It involves two principles:

First, if a model cannot predict the data better than other models, this model should no longer be considered. The model not belonging to set

$$A = \left\{ M_k : \frac{MaxP(M_l|D)}{P(M_k|D)} > C \right\} \tag{2.58}$$

should not be considered. From (2.58), this is just a posterior odds ratio for the model. Model M_k should be excluded since it cannot predict better than other models having higher posterior probability supported by data. This is also similar to George (1995), who uses the posterior odds ratio to approximate the Bayes factor. The constant C is chosen by an appeal to Occam's razor, which excludes models that receive less support from data than submodels (that is, a smaller model nested within the larger model). The set containing

$$B = \left\{ M_k \ni M_1 \in A, M_1 \subset M_k, \frac{P(M_1|D)}{P(M_k|D)} > 1 \right\} \tag{2.59}$$

Models should not be included in (2.44). From these two principles of Occam's windows, (2.44) can be formed

$$P(I|D) = \frac{\sum_{M_k \in A} P(I|M_k, D) P(M_k|D) P(M_k)}{\sum_{M_k \in A} P(M_k|D) P(M_k)} \tag{2.60}$$

where A contains proper models with $A = A' \setminus B \in M$.

The above procedure will reduce the set needed to sum up in (2.44), but this procedure still needs to identify models in A. Two steps carry out the identification, and the first is by Occam's windows. By the posterior odds ratio for the two models, $P(M_0|D) / P(M_1|D)$, where M_0 has fewer independent variables than M_1. The idea shown is if there is strong support for M_0, then M_1 is excluded; but to reject M_0, strong evidence is needed for the larger model, M_1. If the posterior odds ratio falls between (L, R), then neither model is excluded. To determine the endpoint of the interval, Madigan et al. (1994) suggest L = 1 and R = 1/20. The second principle to identify the set of potential models is simple: if a model is excluded, then all other models nested within it are excluded. These two steps identify the model in A; the algorithm follows.

Madigan et al. (1994) provide two algorithms to carry out Occam's windows: the up and down algorithms. For the down algorithm, start with one larger model and subtract variables from the model. For the up algorithm, start with a simple model and add variables to the model. The order of this execution will have some effect on the final set of models.

Now let A and C be subsets of model space M, where A is the set of acceptable models and C is the set of possible models. Begin with $A = 0$ (null model), and C as the set of starting models. Let L and R be the left and right bounds for Occam's windows.

Down Algorithm

1. Select a model M from C

2. $C \leftarrow C - M$ and $A \leftarrow A + M$

this says subtract M from C and add M into A

3. Select a submodel M' from M by subtracting one variable from M

4. Compute the log(posterior odds ratio) = $\log(\frac{P(M'|D)}{P(M|D)})$

5. If log(posterior odds ratio) > R, then $A \leftarrow A - M$ and if $M' \notin C, C \leftarrow C + M'$

this means if the posterior odds ratio is larger than the right bound of the Occam's window, subtract M form A. If the submodel M' does not belong to C then add M' into C.

6. If L $\leq$ log(posterior odds ratio) $\leq$ R, then if $M' \notin C, C \leftarrow C + M'$

7. If there are more submodels in M, go back step 3

8. If $C \neq 0$, go back to step 1

Up Algorithm

1. Select a model M from C

2. $C \leftarrow C - M$ and $A \leftarrow A + M$

3. Select a model M'' by adding one variable into M, that means M'' has one more variable than M.

4. Compute log(posterior odds ratio) = $\log(\frac{P(M|D)}{P(M''|D)})$

5. If log(posterior odds ratio) < L then $A \leftarrow A - M$ and if $M'' \notin C, C \leftarrow C + M''$

6. If L $\leq$ log(posterior odds ratio) $\leq$ R, then if $M'' \notin C, C \leftarrow C + M''$

7. If there are more larger models of M, go back to step 3

8. If $C \neq 0$, go back to step 1

When these algorithm stops, A should contain the potential acceptable models. Finally exclude models which belong to (2.59), and exclude model M_k such that

$$\frac{Max_l P(M_l|D)}{P(M_k|D)} > C \tag{2.61}$$

where C can be chosen as R (Hoeting, 1994), and A contains the acceptable models to be averaged in (2.44). As Hoeting (1994) and Raftery et al. (1997) argue, the number of terms in (2.44) will typically be reduced to fewer than 25 models, and may be as few as 1 or 2. This makes the computation more efficient and less time-consuming.

The second method to carry out the BMA is via the Markov Chain Monte Carlo method, called MC^3 composition derived by Madigan et al. (1995). The MC^3 method evaluates (or approximates) (2.44) directly instead of indirectly using Occam's window. The MC^3 method generates a stochastic sequence that moves through the model space. Let the model space, which contains possible models, be M. Then construct a Markov Chain $\{M(t) = 1,2,\cdots\cdots\}$ with the state space M and equilibrium distribution $P(M_i|D)$. Now generate (or simulate) this Markov Chain for which t = $1,2,3,\cdots\cdots,T$; then under the regularity condition of Markov Chain theory, for function $C(M_i)$ defined on M, the average

$$C = \frac{1}{T}\sum_{i=1}^{T} C(M(t)) \tag{2.62}$$

is a simulation-consistent estimate of $E(C(M))$ (Smith & Robert, 1993). In this research, letting $E(C(M)) = P(I|M,D)$ simplifies computations.

To construct this Markov Chain, it is necessary to define a neighborhood about M, $Nbd(M)$. $Nbd(M)$ contains the set of models with one or fewer models than M. Define a transition matrix p with $p(M \to M')$ for all $M' \notin Nbd(M)$ and $p(M \to M')$ = constant for all $M' \notin Nbd(M)$. If the chain is currently in state M, draw M' from $p(M \to M')$. This is then accepted with probability

$$\min\left[1, \frac{P(M'|D)}{P(M|D)}\right] \qquad \text{(2.63) (Hoeting, 1994)}$$

or stay in M. Raftery et al. (1994) and George (1995) consider using Bayes factor to approximate $\frac{P(M'|D)}{P(M|D)}$. Madigan et al. (1994) adopt MC^3 in their discrete graphical model.

Three of above four methods (MBVS, Geweke and BMA) are used in this research. The Bayesian variable and model selection approach is used to determine appropriate subsets of variables to be used in models of city and metropolitan area growth rates. The outcome of these procedures is compared to those obtained using classical variable (or model) selection techniques.

CHAPTER THREE

GROWTH IN CITIES: URBAN ECONOMICS PERSPECTIVES

The goal of this research is to investigate the sources of city growth using the Bayesian approaches developed in Chapter Two. In economics, many factors must be considered before making any decisions based on models for which uncertainty exists about their exact specification. The question is how do we to select the appropriate models and variables to explain the situation we face? This problem will become more important for economists and decision-makers as the statisticians collect information on more economic variables.

One interesting topic in urban economics is how to explain the growth of cities. Do the cities have any influence on the region or even more on the nation? Jacobs (1984, p. 106) states that the capital that cities generate reaches to remote regions. Jacobs (1969 & 1984) also argues that major cities all over the world can extend their influence by exports to in other ways. Historically, she finds that cities have large effect on the national economy.

Major cities in the U.S. should make some contributions to the national economy as well as to those of other countries. Some cities, like Dallas, Seattle, and Phoenix, have grown very rapidly during the last twenty years and become the major capital contributors for their respective regions (or even more, for the U.S.). Why have these cities grown so fast while others have not? What are the major stimulants of the growth of cities? Given the large number of possible determinants of city growth, how can we select the appropriate factors that explain it? Models and variable selection methods using the classical approach or the Bayesian approach may solve these questions.

The motivation for applying variables and model selection to urban economics came from a paper by Glaeserm, Scheinkman, and Schleifer (GSS, 1995). In this paper, the authors examined the relationship between many urban characteristics and urban growth between 1960 and 1990. GSS find that some characteristics do have a relationship to urban growth, such as human capital, race, and unemployment rates. GSS explain the relationship between those urban characteristics and urban growth in many aspects, for example city growth against manufacturing, unemployment, education, race, government expenditure, region, and income distribution. The authors present many, sometimes conflicting results. One goal of this dissertation is to resolve some of these conflicts using Bayesian analysis. GSS find that some variables (the urban characteristics) are important factors to urban growth but some are not. And this is the motivation for my research, I intend to implement the BVS and BMS into this situation. Using BVS and BMS methods, a statistically coherent means of determining which variables to include or which model(s) to use can be obtained and can explain the relationship between urban growth and various urban characteristics more appropriately. Next I will discuss the variables that will be considered in this application.

3.1 Population

The population for U.S. cities has changed rapidly during the last twenty years, especially as cities have been dispersed to suburban areas. Using population growth as the dependent variable (as measurement for the urban economic growth), GSS try to explain the relationship between growth rate and many city characteristics. Positive population growth may indicate that this city attracted more labor, or more firms, or had better living conditions for workers. This kind of city will attract more people, and these stimulate growth. Within the U.S., labor mobility is affected most by the opportunistic workers have. As labor moves in (or immigrates), productivity should increase and the city grows faster. Although labor immigrates to urban areas, the central parts of cities grow more slowly than suburban areas. Mills and Lubuele (1997) argue that labor tends to emigrate from central city to suburban areas, because of social problems of the inner-city residents. Voith (1998) also examines the relationship between city and suburban growth and argues with that city growth has an effect on suburban growth. Voith's finding contrasts with that of many economists who think that suburban growth is independent of central city growth. Voith also used population growth as a measure of economic growth (along with employment and income growth). As Mills et al. (1997) pointed out that workers move from the central city to suburban areas because of transportation costs, poor living environment, racial problems and poverty. Mieszkowski and Mills (1993) also list possible reasons for resident immigration to suburban areas. These factors reduce the city population but increase the suburban population. As in GSS, I will also consider the population of the Standard Metropolitan Statistical Area (SMSA), or Metropolitan Statistical Area (MSA) as a dependent variable. Using two measurements I will try to reveal which urban characteristics best explain economic growth. The next section discusses the urban characteristics that will be considered as independent variables in my research.

3.2 Education

According to research in urban growth, human capital is one of the most important factors. Lucas (1988) argues that human capital can have an important effect on productivity. Workers who have more experience or knowledge will generate an external effect, or knowledge spillover, to other workers. This spillover effect stimulates worker productivity and also city growth. Simon (1996) uses data from English cities to show that the cities with higher human capital and information grow faster than the cities that have less human capital and information. Rauch (1993) argues that the average level of human capital affects productivity indirectly through the effect of sharing ideas for technological innovation. Simon (1998) believes that cities with a high concentration of highly educated workers should become more productive and attract more people. Simmon also finds a positive relationship between human capital and MSA growth from 1940 to 1986. Glaeser, Kallal, Scheinkman, and Schleifer (1992) use data set from 170 cities in the U.S. to confirm the knowledge spillover effect on urban employment growth.

There is widespread agreement about the importance of human capital on urban growth. To measure human capital, several variables will be considered:

1. Median years of school, which is the average number of years people attend school. This measurement is used by GSS and Rauch (1993) for their analyses. And this is the most commonly used measurement of the human capital.
2. Percentage of population over 25 years old who have 12-15 years in school. This is the percentage of the population that has received a high school degree. Large numbers indicate higher levels of human capital.

3. Percentage of the population over 25 years old who have 16 years or more in school. This is the percentage of the population who has a college or higher degree.

Larger values of the three variables are expected to have positive effects on urban growth.

3.3 Social Characteristics

The social structure of city needs to be considered as a factor in determining economic growth. The social characteristics considered will be race, income inequality, and population age. First, I will discuss the race. Nonwhite residents in a city, especially blacks residents are usually in the lower social and income classes in a city. The stereotype for inner city, blacks is low income, low education, and low labor skills. These three images for blacks reflect the current black residents living in inner cities and are considered to be the main reasons preventing them from moving to a better suburban areas (South and Crowder, 1997). South et al. (1997) also find that for inner city residents, the probability that white residents move to a better suburban area is higher than for the black residents. This result comes from research which indicates that cities having large proportions of low income, and poorly educated blacks tend to grow more slowly. The same may also be true for other minority residents (like Hispanic, and Asian), who engage in low skilled jobs and earn less than most white residents in the city. One potential variable in model of city growth, therefore is the percentage of non-white population as an explanatory variable to see it really has a negative effect on city growth. Based on previous research, the coefficient is expected to be 0 or negative.

Another issue regarding race is the segregation problem in cities. Although many cities in the U.S. are racially integrated (many ethnic groups living together), several of them such as New York, Chicago, and Los Angeles still have segregated areas with very little racial diversity. This segregation has many causes and also may have disadvantages (Kempen & Ozuekren, 1998):

1. The group that has been segregated may not have the information required to get a better job.
2. Children living in a segregated (or ethnic concentration) area will have less chance to receive a good education, especially if they do not speak English well.
3. The negative image of segregation will slow the immigration to the city.
4. Residents who live in the segregated area do not have equaled access to quality social care. This makes the living quality in this area even worse.

The disadvantages of segregation are easy to understand. In this research, I will include this segregation into the analysis to see if it has a negative effect on city growth. The measurement for segregation is adopted from Taeuber and Taeuber (1965) and is called the "segregation index." This segregation index measures how integrated the city is, based on race. If this index is equal to 1, it means this city has only nonwhite (or white) in some areas of the city. If the index is equal to zero, it means this city is perfectly integrated (nonwhite and white residents living together). The other two researches regarding the segregation are Culter and Glaeser (1997) and Culter and Vigdor (1999) which provide good analysis and segregation index (which I will adapt in Chapter Four).

The income distribution in a city is also an important indicator of the economy. If the difference between high income and low-income residents is too large, then it may have some effect (negative or positive) on city growth. This difference probably will make more social problems (especially crime) and slow the city growth. The variables I will use are the following:

1. Percentage of population with income less than $3,000 a year. (The low income class in the city)
2. Percentage of population with income more than $25,000 a year. (The high income class in the city)
3. Per capita money income. This is the total income of the city divided by the population of the city.

These three income-related variables are expected to have some effect on city growth, either positive or negative.

The last variable in this social category is the age of the population. For some cities in the U.S. (or other countries), young people move to other cities and leave an older population in the city. Does this reduce the productivity of the city? That is the reason I have included this variable (the aging population) into this research. These is not much research related this variable to city growth, and not many researchers take it seriously. So I will take the aging population into account, and try to explain its relationship to city growth. The measurement I will use is the percentage of population who is over 65 years old. As the percentage of older people increases, is it good or bad for city growth? This will be discussed in the next chapter.

3.4 Unemployment and Manufacturing Sector

High unemployment within a city is a serious problem because unemployed workers represent an opportunity cost for the city (McDonald, 1997). This opportunity cost of production may cause the city to grow at a lower rate in the long run. GSS (1995) found that the initial unemployment rate reduces economic growth for both city and SMSA. The unemployment rate reduces city growth because of these two effects (GSS, 1995):

1. Workers move to other areas because the high unemployment rate causes business cycle shock; this reduce population growth substantially.
2. Unemployed workers may have skills and professional training, but cannot join in the labor force to stimulate the city growth.

So higher unemployment rates are predicted to slow city growth.

For the production side, the industry sector also plays an important role in urban economics, especially the manufacturing, which employs the largest proportion of workers in urban areas during the last few decades (Mills and Hamilton, 1994). Although other sectors are growing faster then manufacturing sector grows slower; the importance of the manufacturing sector cannot be ignored. In Detroit for example, the automobile industry employs a large proportion of the labor force. In Pittsburgh, the steel industry is the largest employer; many examples follow. Manufacturing is generally thought to induce large multiple effects in the economy, the reason is

First, the manufacturing (automobile, steel, etc.) firms locate in a city (where the firms think they have advantages), then firms employ workers locally (or from other areas) and stimulate the employment rate for this area. Next, firms grow at a fast rate and continue employing workers; this expansion attracts more labor moving to this city, which causes the city population to increase. So this consequence for the manufacturing sector should stimulate city growth. But GSS's empirical evidence indicates that as employment rate of the manufacturing sector increases, city growth rates diminish (e.g. as measured by population growth, manufacturing employment, SMSA population growth and city income growth). Basically, too many manufacturing firms located in one city cause many problems, like pollution, traffic congestion, and social problems. These problems make the quality of life worse and slow immigration into the city, finally reducing the city growth (this reason is probably one of many reasons for the manufacturing firms to slow the city growth). I will use the proportion of employment rate in the manufacturing sector (the labor force in the manufacturing sector) in the city as the measurement, and try to interpret its implication in the empirical studies.

3.5 Geographical Factor

During the last twenty years, many new cities have grown faster than the old cities and they (the new cities) are located in some particular region that have many advantages over other cities, such as a good quality of life, transportation cost is low, and many employment opportunities. As these new cities become more attractive to labor (low skilled or highly skilled), workers start to immigrate to those new cities and this makes these new cities grow very fast. But those cities without as many advantages as fast-growing new cities start to decline. Krugman (1991) argues that this happens because manufacturing becomes concentrated in a few regions but leaves other regions undeveloped. Due to the economics of scale, manufacturing firms will only locate in cities which have the following advantages: the demand for the product is large and the transportation cost for the product is relatively low (Krugman, 1991). This advantage can be confirmed historically, as society spends a large proportion of income on nonagricultural goods; the region with a large population will attract more producers. Then they will mass-produce and economics of scale are formed, because the transportation cost (like railroad and airlines) is lower and the demand in the local market is high. This process will continue and force the traditional agricultural population to concentrate at some region (like the central U.S. states) (Krugman, 1991).

So the main argument for Krugman is that there is a diverging trend by region in the national economy. Some cities grow faster than some undeveloped cities in the nation. Krugman (1991) also develops a two-region model with two types of production: agriculture and manufacturing goods. Krugman found that the region with the lower transportation cost and the higher manufacturing share (large economics of scale), will attract more manufacturing firms, which will make this region grow faster than the other region.

Barro and Sala-Martin (1992) provide a different argument. Using 48 U.S. states and 98 countries as data, they found there is evidence for convergence. This means that poor states (or countries) grow faster than rich states (countries) in terms of per capita personal income. Barro et al. (1992) used a neoclassical model set-up to do the empirical studies, and found the economics tend to grow faster in per capita terms when they are far below the steady-state position (which is clear for the 48 U.S. states from 1840 to 1988) (Barro et al., 1992). But the main difference for Barro's measurement is the population growth. So the region effect needs to be specified in the empirical part of this research. Despite the divergence or convergence point of view (Krugman and Barro et al.), the geography factor is an important factor in economic growth and this is the reason I will include a geographical factor into this research. The classification for each city will be based on Tauber (1965), and I will consist of regional dummy variables. I expect there to be regional difference in economic growth.

3.6 Government

There are many arguments about the government's role in city growth. Steinacker (1998) argues that the force of economic restructuring and de-industrialization are major concerns for local government. But many of the factors cannot be controlled by local government; some non-metropolitan areas need to be considered also. Bradbury, Downs, and Small (1982) argue that policy intervention could stop the decline of American cities, which implies that the local government needs good policies to break the negative feedback cycle of city decline. For Bradbury et al. (1982), policies can correct the local market failure only through appropriate adjustment. So only "good" policies are good for urban growth. This (consequences of policies) reflects the importance of government, because government is the executor for policy and take responsibility for it. How does government affects the city growth and by what channel? The channels I will consider are those of GSS, which are the government expenditure in various categories and government revenue, as follows:

1. Education: Government (either state or local) has spent the largest share of its budget on education during the last 30 years, although the budget share has declined in recent years. The main reason for this decline is that the school-age population is mainly spent on higher education (college and universities); which trains highly skilled labor for cities, and this source has constituted the

main resource of human capital. Human capital, as described previously, is one of the most important sources of urban growth. So as the percentage of expenditure on education increases, it should have an impact on urban growth.

2. Transportation: the transportation system of a city is very important; not only does it provide the access for residents to the city, it also contributes to provide the quality of life for the city. The main transportation system in cities is the highway system, which is financed either locally or at the state level. Voith (1993) argues that the value of highway accessibility (to the city center) parallels the economic performance of the city. Voith (1993) also argues this accessibility makes the inner city grow fast, and benefits the surrounding suburban area. So the main argument for Voith is that the suburban areas and the inner city cannot be isolated; in fact they need to co-exist and will benefit each other.

   The link between these two areas is the highway system, which not only provides accessibility to each area, but also stimulates growth for both areas. So the importance of the transportation system is clear. The role of government is to build a good transportation system for a city, and this can be executed through the transportation system budget. So I use the percentage of expenditure for highways as a measurement, and see whether it has an effect or not.

3. Public Safety: This category has become a major concern in cities recently, especially in inner cities. The crime problem associated with the inner city can reduce growth substantially and make the quality of life worse. Mills et

al. (1997) argued that the crime problem is serious in the U.S. and the government needs to focus on it. Trillions of dollars have been spent to fight crime, like drugs, but the effect is not good. The main reason is that policy has not been designed well or administered well, so policy change may improve the situation (Mills et al., 1997). Although good policy is needed, the budget for public safety protection still provides a good measurement.

4. Health Care: Cities with better medical facilities and health care should attract more labor; hence they are good for city growth. The city government can provide quality health care via building more modern medical facilities (like hospitals), which can provide services for city residents. So the government expenditure health care (hospitals) should reflect how important the government deem this area.

5. Sanitation: Water and sewer service are among the most important elements for a good living environment. Water is the most important factor for human health and whether it is clean or not will affect health directly, the same as the sewer an garbage disposal; proper disposal could prevent epidemics and possible environmental pollution. So the government should consider that this will be very important in the future, and increase the budget for the sanitation service because it is very important to increase the quality of living for city residents.

6. Government Revenue: This is the fund that the government raises to finance its operation. The sources I will consider are the following: tax,

intergovernmental transfers, debt, and general revenue of the local government.

General revenue is the total amount of funds collected or raised by the local government. It comes from many sources; taxes (income tax, property tax, sales tax, and corporate income tax), utility revenues, other miscellaneous charges and revenue from other levels of government (state or federal). As the city grows, there should be many sources of revenue moving in, which gives the local (city) government more funds to put in the public sector (such as building more convenient transportation systems and preventing air and water pollution). This (increasing of revenue) will have a positive effect on the city growth and benefit the surrounding area. So I expect fast growth in cities to be associated with high revenues.

The main source of revenue for local government if from various taxes collected from residents; sales tax, property tax and corporate tax. Property tax (especially real estate tax) is the main tax source for local governments, contributing nearly 40% of local revenue, but it has decreased during recent years. The source I will use is the percentage of government revenue from tax, which includes all taxes collected by the government. This is the main financial resource for the government (approximately 50% from Table 2) to collect and should have effect on city growth, either positive or negatively.

Intergovernmental transfer to finance local government has become more important in recent years. The main reason is that the state (or federal) government thinks the city with a growing population needs more funds to operate. It can be confirmed that the city is deemed by state and federal government as an important area to develop. As a city grows faster, intergovernmental grants from the state have become the largest source of funds for local governments in the U.S. (McDonald, 1997). This is why I will consider it as a factor that affects the city growth.

The last candidate I will consider is the debt raised by government. Debt is a way for governments to raise funds by borrowing from either corporation or residents. The return may be as interest, share of government-operated business, or another form. If the resident or corporation thinks the city has a higher growth potential, they may be willing to help (or finance) local government by buying bonds which helps the government to expand. But whether it is positively correlated with city growth or not is an empirical matter and will be investigated in the next chapter.

The data used to analyze include two city files: one is the central city data referred as CITY (77 cities), and the other is the MSA data referred as MSA (75 MSA). CITY contains 23 variables while MSA contains 21 variables. There are some differences between these two data sets (some data are available for the CITY but not for MSA), which will be described in Appendix A in detail. The segregation index (measured by the dissimilarity index) is not available for all 77 cities in the CITY and 75 cities in the MSA. Only 63 cities and MSA have a dissimilarity index available. I use the segregation index in the social category which is in Appendix B. The cities that do not have this particular index are listed in Appendix A.

CHAPTER FOUR

MODEL AND VARIABLE SELECTION APPLICATION IN URBAN ECONOMICS

In this chapter, I implement a classical and Bayesian approach to the model and variable selection problem. First I use the usual regression analysis to examine the relationship between the economic growth and each city's characteristics. As in GSS, I divide the data into different categories, for example, growth and education or employment, growth and government factors, etc. After these categories have been analyzed, I put all city characteristics into the analysis to analyze their relationship to measure of city growth. Then I implement the classical approach for variable or model selection to see what the classical approach explains. In next stage, the Bayesian approach is implemented. These are the three approaches discuss in Chapter Two and all the selection results are listed in Appendix B (Table of the Analysis). For all the variables used under GSS's category (from Table 4 to Table 10), are listed in Appendix A.

4.1 Prior Setting for Bayesian Approaches

Before implementing the Bayesian approach, I will briefly describe the prior setting for each of the three methods implemented in this dissertation. For the BMA, the prior for the parameters is set as $N(m, \sigma^2 V)$ where m is set as $(b_0, 0, 0 - - - 0)$, b_0 is the sample mean of dependent variable (Y), and V is set as (2.48). And $v\lambda / \sigma^2 ; \chi^2_v$ where $v = 4.00, \lambda = 0.25$ and $\Phi = 3$ is set as the default value in Matlab program for the hyperparameters. For Geweke's approach, prior probability that each variable appear in the model is set as 0.5 and prior precision is set as (2.29) which is referred to Adkins et al.(1999) for detail. For the MBVS, r is set as $0_{1\times k}$ where no variables are included before sampling; $\delta = 3$ as stated in Brown et al. (1998) and $w = 0.5$ as Geweke's prior Bernoulli probability to include variables. Other settings for the prior are described in Chapter Two. For Geweke's approach, the P.M.P. means posterior marginal probability (the posterior probability that the coefficient = 0) for convenience. And M.P.P. refers the model posterior probability for these three methods for convenience.

4.2 The Regression Analysis for City Growth

In this section a traditional regression analysis is conducted to gain a basic understanding of the relationships under study. Two data sets are used: First is referred to as the central city data set (CITY) and second as the MSA data set (MSA). The dependent variable for the CITY and MSA data sets is the population growth of the central city and the MSA. From the summary statistics in Table 1 and Table 2 one can see that the mean population growth for the MSA (0.406) is higher than that for the central city (0.115). This supports one general fact: The MSA population has been increasing during the last twenty years, because of the increasing population in the suburban area. For education, the MSA has a higher percentage of both high school and college graduates (57.42% to 54.57% and 12.10% to 12.05%) compared to the central city (CITY). Also the MSA has a lower unemployment rate than the central city, which is 4.33% for the MSA and 4.77% for the CITY. The central city also has a higher percentage of both non-white population and aging population (which I define as percentage of the population that is more than 65 years old) than the MSA. This suggests another fact: the labor force is moving toward the suburban areas rather than to the central city. This may explain why the MSA population growth is higher than that in the central city and could be one of the factors used to explain the differences between the MSA and the CITY. Next, the relationship between the population growth (which is used to measure the economic growth) and a variety of city characteristics is analyzed as in GSS Tables 4 to 10. In addition, I will also analyze city growth using all city characteristics, which were not examined in the GSS study.

4.3 City Growth and Manufacturing (GSS Table 4)

First, I use the traditional approaches, like R^2, AIC, BIC and SBC to select models. From Table 3 in Appendix B, the model that has the lowest AIC and BIC selected models includes Lpop70, Lpc70, Mfgs70, South, Central and NEast. Central and NEast are in all six models that are selected. The similar results for the MSA data; the model includes Lpop70, Central and NEast. Again, Central and NEast are in all six models that are selected; in addition Lpop70 is in every model. So from the traditional approach (according to AIC and BIC) results, Central and NEast are the most selected variables, and Lpop70 also appears to be important.

Second is the BMA approach, which uses the MCMC method to obtain results. From Table 5 (for the CITY), the model that has the highest posterior probability includes South, NEast, Lpc70 and Lpop70. For this BMA procedure (for both city and MSA), I drew 10,000 Gibbs samples (the burn in samples are set as 10% of the total samples for the BMA) with 7 independent variables. For the MSA, from Table 6, the model that has the highest posterior probability includes South and NEast. From BMA, South is the most selected variable for both CITY and MSA. It also has the lowest posterior t-probability (0.0004 for CITY, 0.08 for MSA.) of the variables for the CITY and MSA. The posterior t-probability, like the usual t-probability, is the probability the coefficient which is equal to zero based on the posterior t-statistics. Lpc70 has a lower posterior t-probability (0.009 for CITY, 0.15 for MSA) than Lpop70 for both CITY and MSA. Central is the least selected variable for these categories, which has a higher posterior t-probability than other variables (0.883 for CITY, 0.845 for the MSA). So in the BMA approach, the models selected include the variables that have the lower posterior t-probabilities (in most cases) and exclude the variables that have a higher posterior t-probability.

Third is Gweke's approach, which uses the subjective prior (the prior probability for each variable to be included is set as 0.5) and also uses MCMC to carry out the computation. For this, the number of Gibbs samples drawn is 10,000 and the first 2000 samples are discarded for the burn-in. From Table 7, the model that has the highest posterior probability (= 0.4950) includes South, Central, NEast, Lpc70 and Lpop70 for the CITY. For the MSA (Table 8), the model that has the highest posterior probability (= 0.2383) includes Central, NEast, Mfgs70 and Lpc70. Central, NEast, Mfgs70 and Lpc70 are the most selected variables. For this category, Geweke's approach selects the same variables (including Central, NEast, Mfgs70 and Lpc70) which has a higher posterior probability.

Fourth is the MVBS approach, which is similar to the SVSS approach. It has different prior setting but is easy to implement. For this MBVS approach, 10,000 Gibbs samples are drawn (with 10% samples for burn in) and prior settings are as stated in **4.2**. The model (Table 9) for the CITY that has the highest posterior probability (= 0.1095) includes Central, NEast, and Mfgs70. For the MSA (Table 10), the model (posterior probability = 0.0689) includes Central and NEast. The most frequently selected variables for the CITY are Mfgs70, Central and NEast. Lpc70 is selected in several models but not more than the three variables stated above. The most selected variables for MSA are Central, NEast and Lpop70.

From the results above see that, different approaches produce different models due to different prior settings (except AIC and BIC), the geographical factors are the most selected variables in this category, same as in GSS which measured in term of t-statistics (city growth and manufacturing). The employment percentage in manufacturing is not critical for the MSA but is probably critical for the CITY. This result is similar to GSS for CITY but not for the MSA, which GSS states as an important variable.

4.4 City Growth and Unemployment (GSS Table 5)

This section explores the relationship between city growth and unemployment. For the traditional approach (Table 11), the model that has the lowest AIC and BIC includes Lpop70, Mfgs70, South, Central and NEast for the CITY (AIC = -206.328 and BIC = -203.082). For the MSA (Table 12), the model includes Lpop70, Central and NEast (AIC = -182.767 and BIC = -180.179). All five models for the CITY select South, Central, NEast and Mfgs70. Unemployment is only selected in two CITY models, and less often than the geographical factors. All five models select Lpop70 and Central and four models select NEast for the MSA. Unemployment is only selected by one model, which also includes Lpop70, Central, and NEast. The results of the traditional approach show that the geographical variable is the most important factor in this category but unemployment is not.

For BMA, from Table 13 of the CITY, the model that has highest posterior probability includes South, Mfgs70, Lpc70, and Lpop70 (posterior probability = 0.15161). South and Lpc70 are selected by all six models in Table 13 with posterior t-probability 0.000421 and 0.015511, respectively. Also the other frequently selected variables for the city is Lpop70 (posterior t-probability = 0.020856) which is selected by five models in Table 13. Unemployment is only selected by one models in Table 20, with a posterior t-probability = 0.707. For the MSA, the model that has the highest posterior probability includes South, Lur70, Lpc70, and Lpop70 (posterior t-probability = 0.06416). Central is selected by four models in Table 14, with posterior t-probability 0.092417. Lur70 (posterior t-probability = 0.7070) is selected by three models in Table 14, which contrasts with the results from Table 13, where Lur70 is only selected by one model. As for the MSA, Lpc70 is also selected by three models in Table 14 with a posterior t-probability of 0.093092. The above results that the geographical variables may still be major factors to consider in this category. Unemployment is important for the MSA but probably not important for the city.

For Geweke's approach of the CITY (Table 15), the model that has the highest posterior probability contains South, Central, NEast, Mfgs70, Lpc70, and Lpop70 with posterior probability 0.3287. South, Central, NEast, Mfgs70, and Lpc70 are selected by all four models in Table 15. Unemployment enters in two models in Table 15, the same as Lpop70. In Table 15, one model selects all seven variables in this category to explain the relationship between unemployment and city growth. Although this model has lower posterior probability (0.148) compared to the other models, it still can provide a good explanation for the city growth. For the MSA, the model with the highest posterior probability contains Central, NEst, Mfgs70, and Lpc70 with a posterior probability of 0.140. Central, NEast, and Lpc70 are selected by all five models in Table 16. Unemployment is selected by one model in Table 16, less than the CITY. But the posterior marginal probability of Lur70 is 0.6919 for CITY and is 0.6055 for MSA (which are closed). This posterior marginal probability states that Lur70 should not be selected frequently (69.19% for CITY and 60.55% for MSA) and the result is acceptable (is selected by two models for CITY and selected by one model for MSA). But with the flexiablity of Geweke's approach, we can force a higher prior probability that Lur70 enters the model if desired. For the MSA, no models contain all seven variables as in the Table 15. The region variables are still important and Mfgs70 enters three models (in Table 15, Mfgs70 enter all models).

For the MBVS (Table 17 and Table 18), 10,000 Gibbs samples were drawn and the CITY model with the highest posterior probability selects Central, NEast, and Lpc70 (with probability 0.0716), which is the second most selected model of MSA. Central and NEast are the most selected variables (enter all five models in Table 17). Unemployment is included in two of the four models in Table 17, which is similar to the Geweke's estimates. For the MSA data (Table 18), the model with the highest posterior probability selected Central and NEast (with probability 0.0389). The models selected are similar to those in Table 17 for the CITY, but posterior probability of each is much smaller. All six models selected Central and NEast in Table 18, same as Table 17. Unemployment enters into two models (posterior probability = 0.0336 and 0.0329, respectively) in Table 18, same as Table 17.

From the above method, the region dummy variables are still important factors to consider (same as in GSS). The demographic variable (initial population, Lpop70) is also selected frequently (also same as GSS). The initial economic index (initial per-capita money income) is another important variable in the model. GSS concludes that unemployment is an important variable, but it is not selected frequently in this category by any of the Bayesian methods.

4.5 City Growth and Education (GSS Table 6)

Many researches find that education is an important factor affecting economic growth. This section discusses the relationship between city growth and education variables. From Table 19 of the CITY, the model that has the lowest AIC and BIC contains Central, NEast, and High70 (AIC = -217.8 and BIC = -214.7). Central, NEast, and High70 appear in the top five models (Table 19), while Coll70 only enters into one model (with Lpc70, Central, and Neast). For the MSA (Table 20), the model having the lowest AIC and BIC has Central, NEast, Mfgs70 and Lpmed70 (AIC = -183.2 and BIC = -180.1). Central enters all five models in Table 20, which is the same as Table 19 of the city. NEast and Lpmed70 have been selected in four models in Table 20. Lpmed70, weighted by the population of 1970 with Medsy70, has also been selected in four models but is highly correlated with Medsy70. So, one should be cautious in using it in the same model as Medsy70.

For the BMA, from Table 21 of the CITY, the model that has the highest posterior probability selects South, NEast, Coll70, Mfgs70, Lpop70 and Lpmed70 (with posterior probability = 0.0327). South (with posterior t-probability = 0.002405), Lpop70 (with posterior t-probability = 0.000098) and Lpmed70 (with posterior t-probability = 0.1484) are selected by all six models in Table 21. Coll70, the indicator for a higher human capital level, is selected in five models with posterior t-probability of 0.19033. The other indicator for human capital, Medsy70, with a very high posterior t-probability of 0.9751, is not selected by any model in Table 21. Unemployment is selected in one model but with high posterior t-probability of 0.752. For the MSA, from Table 22, the model that has the highest posterior probability selects High70, Lur70, Mfgs70 and Lpmed70 (posterior probability = 0.13499). High70 is selected in all five models in Table 22 with posterior t-probability = 0.0323. Lur70 is selected by three five models with posterior t-probability = 0.2337, which is lower than in CITY. Coll70 has a low posterior t-probability in CITY than in MSA. High70 has low posterior t-probability in the MSA data, but has a high posterior t-probability for CITY.

For Geweke's approach, from Table 23 of the CITY, the model (posterior probability = 0.1000) that has the highest posterior probability has Central, NEast, High70, Lpc70, and Lpop70. Central, NEast, High70 (P.M.P. = 0.0000), and Lpc70 (P.M.P. = 0.1855) enter all seven models in Table 23. Coll70 (P.M.P. = 0.7941) has not entered any model and Medsy70 (P.M.P. = 0.4836), enters into three models. For the MSA, from Table 24, the model (posterior probability = 0.0491) with the highest posterior probability has Central, NEast, Medsy70 (P.M.P. = 0.2849), Lpc70 (P.M.P. = 0.253), and Lpop70. Central and NEast are selected by all seven models in Table 24. Four models in Table 24 select Lur70 (P.M.P. = 0.5611) and six models select Lpc70. The P.M.P. (which referred as posterior marginal probability that coefficient = 0, or the probability that variables should be omitted) for the education variables are as follows: For CITY, High70 = 0.0000, Coll70 = 0.7941, and Medsy70 = 0.4836. For MSA, High70 = 0.7335, Coll70 = 0.8224, and Medsy70 = 0.2849.

For MBVS, from Table 25 of the CITY, the model has the highest posterior probability (= 0.0607) selects Central, NEast and Medsy70. All five models in Table 25 select Central and NEast. No education variables are selected in Table 25, which contrast with previous two Bayesian methods. From Table 26, the model that contains Central and NEast has the highest posterior probability (= 0.0243). All six models in Table 26 select Central and five models select NEast. No model selects education variables in Table 26 (same as Table 25) and Lur70 is selected by two models. Adkins et al. (1999) find that the results proceeded by Geweke's procedure is relatively sensitive to the chosen prior information and conclude that good results depends on choosing a 'good' prior.

From the three Bayesian approaches, the region variables (especially Central and NEast) are still important variables to consider. The several education variables, like Medsy70, Coll70, or High70, are also important due to the different methods of selecting variables, but they are not as important as region variables (which is very similar to the results of GSS). Unemployment enters several models and these models could give a good explanation about its relationship with education and city growth.

4.6 City Growth and Inequality (GSS Table 8)

This section analyzes the relationship between income inequality and city growth. From Table 35 of the CITY, the model that contains Central, NEast, Lpop70, Lpc70, Mfgs70, Incle70 and Lmedic70 has the lowest AIC (= -223.8) and BIC (= -219.3).

Central, Lpop70, Mfgs70 and Incle70 are selected by all models in Table 35. For the MSA, from Table 36, the model that contains Central, Lpop70 and NEast has the lowest AIC (= -182,8) and BIC (= -180.2). All six models in Table 36 select central and Lpop70, which is the same as Table 35. No model in Table 36 selects Incle70, in contrast to Table 35.

For BMA, from Table 37, the model (P.M.P = 0.0999) with South, Mfgs70, Medsy70, Incle70, Lpop70, Edle70, and Lmedic70 has the highest posterior probability. South (posterior t-probability = 0.0542), Mfgs70 (posterior t-probability = 0.0150), Edle70 (posterior t-probability = 0.000016) and Lmedic70 (posterior t-probability = 0.000666) enter every model in Table 37. Incle70, one important factor when considering income inequality, has entered five models with posterior t-probability = 0.0817. For the MSA, from Table 38, the model with South, NEast, Incle70 has highest the posterior probability (= 0.03693). NEast (posterior t-probability = 0.036) are selected by all six models. Lur70 (posterior t-probability = 0.7540) and Incle70 (posterior t-probability = 0.2627) are selected by four models in Table 38. The results from Table 38, although not as fully expected as those from Table 37 (Edle70 is not selected in Table 38), are acceptable. Incle70 and Lur70 (without Edle70) are in the same model to explain the relationship between inequality and MSA growth that is expected.

For Geweke's approach (Table 39), the model that contains Central, NEast, Medsy70, Incle70, and Lpop70 has the highest posterior probability (= 0.581). Central (P.M.P. = 0.0000), NEast (P.M.P = 0.0000), Medsy70 (P.M.P. = 0.0000), and Incle70 (P.M.P. = 0.0000) are selected in ten models in Table 39 (the P.M.P. is consistent with the model selection result). Lur70 (P.M.P. = 0.4652) and Coll70 (P.M.P. = 0.8190) are selected in one model in Table 39. This result from the CITY provides a good explanation between inequality (with Incle70 and Edle70 (P.M.P. =0.3098) entering in the model that has the high posterior probability) and CITY growth, which is as we expect. Medsy70 is also important in this category to help explain the relationship between inequality and city growth. For MSA (Table 40), the model that has the highest posterior probability (= 0.044) selects Central (P.M.P. = 0.0254), NEast (P.M.P. = 0.2209), Medsy70 (P.M.P. = 0.1850), Lur70 (P.M.P. = 0.5281), Lpc70 (P.M.P. = 0.3423), and Lpop70 (P.M.P. = 0.2569). Central, NEast, and Lpc70 are selected by all seven models in Table 40. No model in Table 40 selects the inequality variables (Incle70 (P.M.P. = 0.7951) and Incla70 (P.M.P. = 0.7320)). This result is quite different from that of Table 39, in which the inequality variables (Incle70) is selected by all models. But four models in Table 40 select the education variable (Medsy70, P.M.P. = 0.1850).

For MBVS, Table 41 for the CITY, the model that contains Central, NEast, and Medsy70 has the highest posterior probability (0.0249). All models select Medsy70 and Central, four models in Table 41 select NEast. No model selects the inequality variables and two models select Lur70. For the MSA (Table 42), the model with Central and NEast has the highest posterior probability (= 0.023). All models select central and NEast, two models in Table 42 select Lur70 and Lpc70.

From the results of the Bayesian approaches, the inequality variables are quite important (for BMA and Geweke, but not for MBVS) in explaining city growth. This is consistent with many studies: the higher the proportion of low-income and low-educated population, the slower the city grows. For MSA growth, the regional variables are still the major factors to consider than the inequality variables; the inequality variables are much less likely to be important MSA growth. The inequality variables are not important in GSS (except for Edle70 in CITY), but unemployment and regional variables (same as the Bayesian) are. The Bayesian approach, is more flexible because it allows more variables to explain the relationships between city growth and inequality.

4.7 City Growth and Social Characteristics (GSS Table 9)

For this section, some observations must omitted from the sample since not all 77 cities have a segregation index (which is measured by a dissimilarity index) available. Only 63 cities and MSAs have an available segregation index to use in this section. In addition to GSS Table 9, I add Age70 into this section, in order to see whether an aging population can explain the city growth.

From Table 43, the model that has the lowest AIC (= -261.3) and BIC (= -255.2) contains South, Central, NEast, Mfgs70, Age70, and Lpop70. South, Central, NEast, Age70, Mfgs70, and Lpop70 are selected in all five models in Table 43. Seg70 is only selected by one model and the weighted segregation (Weseg70) is selected by three models in Table 43. For the MSA (Table 44), the model that contains Central, NEast, Nonw70, Mfgs70, Medsy70, Lpc70, and Lpop70 has the lowest AIC (= -159.6) and BIC (= -154.4). All models in Table 44 select Nonw70, Mfgs70, Medsy70, Lpc70 and Lpop70. Seg70 and Weseg70 do not enter into any model in Table 16, which is not as expected. But Nonw70 enters into all models for the MSA and Age70 enters into all models in Table 44, as expected.

For the BMA of this category, 30,000 Gibbs samples are drawn and 3,000 samples are discarded. From Table 45, the model that has South, Central, NEast, Lpc70, Seg70, and Weseg70 has the highest posterior probability (= 0.04063). South (posterior t-probability = 0.000), Central (posterior t-probability = 0.006986), Lpc70 (posterior t-probability = 0.000089), Seg70 (posterior t-probability = 0.045675), Lpop70 (posterior t-probability = 0.0000), and Weseg70 (posterior t-probability = 0.006563) are selected by all six models in Table 45. Age70 and Nonw70 are selected in two models but with high posterior t-probability. For the MSA (Table 46), the model that has South, NEast, Mfgs70, Age70, and Weseg70 has the highest posterior probability (= 0.08722). South (posterior t-probability = 0.13184), NEast (posterior t-probability = 0.003592), Mfgs70 (posterior t-probability = 0.158796), and Age70 (posterior t-probability = 0.027105) are selected by all models in Table 46. In contrast to Table 45, Seg70 and Nonw70 do not enter into any model but Age70 is selected by all models in Table 46. Weseg70 (posterior t-probability = 0.833786) is selected in two models in Table 46 and Lur70 (posterior t-probability = 0.725178) is selected in three models in Table 46.

From Table 47, the model that has the highest posterior probability (= 0.1050) selects South (P.M.P. = 0.0868), Central (P.M.P. = 0.0000), NEast (P.M.P. = 0.0003), Age70 (P.M.P. = 0.0000), Mfgs70 (P.M.P. = 0.3320), Medsy70 (P.M.P. = 0.0000), Lpc70 (P.M.P. = 0.2106), Seg70 (P.M.P. = 0.4222), Lpop70 (P.M.P. = 0.4124), and Weseg70 (P.M.P. = 0.0873). All models in Table 47 select South, Central, NEast, Age70, Medsy70, Lpc70, and Weseg70. Seg70 (in five models) and Weseg70 (in all models) are selected frequently as expected. But Nonw70 (P.M.P. = 0.8093) is not selected in any model, which I did not expect. Now from Table 48 (for the MSA), the model that has the highest posterior probability (= 0.026) selects Central (P.M.P = 0.1878), NEast (P.M.P. = 0.2937), Medsy70 (P.M.P. = 0.2479), Mfgs70 (P.M.P. = 0.1031), Lpc70 (P.M.P. = 0.3334), and Lpop70 (P.M.P. = 0.5475). Central, NEast, Medsy70, and Mfgs70 are selected by all models in Table 48. Seg70 (P.M.P. = 0.6223) is selected in one model (less than in Table 47) and Weseg70 (P.M.P. = 0.8218) is not selected by any model in Table 48. Age70 (P.M.P. = 0.8090) is not selected by any model in Table 48, which in contrast with Table 47, where it is selected by all models.

From Table 49, the four models that have the same highest posterior probability (= 0.0117) selects Central, NEast, Medsy70, Lpc70, and Seg70. Central and Medsy70 are selected by all models in Table 49. Seg70 has entered into three models in Table 49 but no model selects Weseg70. No model select Age70 and Nonw70, not as I expected (also different from BMA and Geweke). From Table 50, the two models that have the highest posterior probability (= 0.0089) select Cental, Lpc70, and Lpop70. Central is the most selected variable in Table 50, being selected by all models. Seg70 is selected by three models in Table 50 but no model selects Weseg70. Like Table 49, Age70 and Nonw70 are not selected in any model.

For this category, various models select the social variables (Seg70, Age70 and Nonw70) to explain city growth. These results confirm many studies that show that these social variables have been associated with low city growth for many years, in particular the non-white population and segregation, which have been considered by many social scientists as the major factors for low central city growth. In this research, I add a variable to measure an aging population into this category and find that various models select it with high a posterior probability. This result confirms my hypothesis that younger people moving to suburban areas, and that an increasing the aging population leads to slower growth in the central city. This aging population could also combine with other social variables (segregation and non-white population) to make the central city even worse. For GSS (they did not put all social variables in the MSA), the social variables are not so important as in Bayesian method. This may result from the Bayesian approaches' (especially for BMA and Geweke) flexiblity to include important variables in the models.

For Table 70, it illustrates the coefficient estimates for three procedures. For GSS, it is the least square estimate. For Geweke's, it is the posterior mean of coefficients and for BMA, it is the posterior estimates of the coefficients. As we can see from Table 71, the coefficient estimates for regional variables are all negative for three procedures. For unemployment rate, it is negative for GSS but is positive for Geweke and BMA (although is not so significant). For % of nonwhite, are all negative for three procedures, but not so significant for three procedures. The segregation index, is negative for GSS (not significant) and BMA (is significant), but is positive for Geweke's. The weighted segregation index (weighted by multiplying % of nonwhite), is positive for GSS (not significant) but is negative for both Geweke's and BMA (is significant). But there are some difference between GSS and two Bayesian procedures. For GSS, the initial year is 1960 but for two Bayesian procedures, the initial year is 1970. For consistent with the data set I used, Table 70 also includes the GSS analysis which the initial year is 1970. The coefficients from Geweke tend to be smaller in magnitude; this occurs because the posterior distributions have significant mass at zero. The MBVS excludes variables more often and the shrinkage of the model averages to zero tends to be greater.

4.8 City Growth and Government (GSS Table 10)

For this section, we are examining the government sector. There are differences between the CITY and the MSA for this category due to the available government information (which will explain in Appendix A). That means that different variables are used in the CITY and the MSA to do the analysis. In Table 51, the model that has the lowest AIC (= -217.3) and BIC (= -212.0) selects NEast, Central, Medsy70, Mfgs70, Expo70, and Ldebt70. For the government variables, only Expo70 and Ldebt70 are selected in four models. Central, Medsy70 and Lpop70 are selected by all five models in Table 51. From Table 52, the model that has the lowest AIC (= -190.8) and BIC (= -185.2) selects Central, NEast, Lur70, Mfgs70, Lgvpc70, Pctax70, Exedu70, Lpc70, and Lpop70. Central, Neast, Lur70, Lgvpc70, and Lpop70 are selected by all models in Table 52. Pctax70, another government variable, is selected by two models.

From Table 53, 30,000 Gibbs samples are drawn (with 3,000 samples are discarded) and the model has the highest posterior probability (= 0.01504) selects South, NEast, Mfgs70, Medsy70, Igr70, Exhwy70, Lpc70, and Lpop70. South (posterior t-probability = 0.247308), NEast (posterior t-probability = 0.006613), Mfgs70 (posterior t-probability = 0.705308), Exhwy70 (posterior t-probability = 0.180879), Lpc70 (posterior t-probability = 0.000248), and Lpop70 (posterior t-probability = 0.123979) are all selected by all four models in Table 53. Other government variables are selected including Igr70 (selected by three models with posterior t-probability 0.668821) and Lpcex70 (selected by one model with posterior t-probability 0.798257). For the MSA (Table 54), the model that has the highest posterior probability (= 0.01568) selects Central, NEast, Medsy70, Mfgs70, Igr70, Pctax70, Lpcex70, and Exhwy70. NEast (posterior t-probability = 0.084475), Mfgs70 (posterior t-probability = 0.135017), Igr70 (posterior t-probability = 0.153024) and Exhwy70 (posterior t-probability = 0.015273) have been entered into all 8 models in Table 34. Other government variables are selected includes Pctax70 (selected by five models with posterior t-probability 0.685314) and Lpcex70 (selected by two models with posterior t-probability 0.770924).

From Table 55 (30,000 Gibbs samples are drawn with 6,000 samples are discarded), the model that has the highest posterior probability (= 0.012) selects South (P.M.P. = 0.3295), Central (P.M.P. = 0.0000), NEast (P.M.P. = 0.0133), Mfgs70 (P.M.P. = 0.0672) Medsy70 (P.M.P. = 0.0000), Lpcex70 (P.M.P. = 0.4925), and Ldebt70 (P.M.P. = 0.1014). All models in Table 55 select South, Central, NEast, Mfgs70, Medsy70, and Ldebt70. Other government variable selected includes Lgvpc70 (P.M.P. = 0.5490) and Expo70 (P.M.P. = 0.5408). But Exhwy70 (P.M.P. = 0.7945) is not selected in any model, which contrasts with the results of the BMA. These selected government variables should provide a useful explanation for city growth. From Table 56, the model that has the highest posterior probability (= 0.022) includes Central (P.M.P. = 0.0134), NEast (P.M.P. = 0.1938), Mfgs70 (P.M.P. = 0.3455), Lpcex70 (P.M.P. = 0.1859), Lpc70 (P.M.P. = 0.2761) and Lpop70 (P.M.P. = 0.1352). Central, NEast, Lpcex70, and Lpop70 are selected by all models in Table 56. For government variables, Lpcex70 (selected by all models) and Lgvpc70 (is selected by one model with P.M.P. = 0.6477) are selected by various models. But as the same as Table 55, Exhwy70 (P.M.P. = 0.8971) is not selected by any model.

From Table 57 (30,000 Gibbs samples are drawn with 3,000 samples are discarded), the model has the highest posterior probability (= 0.0221) selects Central, NEast, and Mfgs70. For Table 57, no government variables are selected which is contrast to previous two Bayesian approaches. As we can see, MBVS selects a smaller model when explanatory variables are larger. From Table 58, the model that has the highest posterior probability (= 0.004) select Central and NEast. Lgvpc70 is selected by two models in Table 58 and is the only government variable being selected. For this category, the CITY and the MSA have 15 explanatory variables, which is a larger number than the previous category. Those models selected by MBVS have fewer variables than the other variable selection methods. This will be confirmed when we put all variables into the analysis.

For this category, various models select government variables but many government variables have not been selected. Exhwy70 should be one important variable since the construction of highways is a major project for any city. When the highway is built, it will bring more business into this area and more labor will move in because the transportation improves. Other government variables, like Igr70 and Ldebt70, represent main financial resources for the local government. When the local government uses these funds efficiently (like building more public schools and improving living quality); it should have a positive effect on the city's economy. For GSS, the government variables are not important, which will not be selected because the t-statistics is very small. For Bayesian approaches, although not every government variables are selected, still includes various government variables (especially for BMA and Geweke) to explain their relationship with city growth.

4.9 City Growth and All Variables (Without Segregation)

In the last part of this Chapter, all of the variables are included into the analysis to see which variables are selected. For the Bayesian approaches, for each method (for the CITY and the MSA) we draw 50,000 Gibbs samples (with some samples will be discarded for burn in) because the number of explanatory variables is relatively large (23 for the CITY, 22 for the MSA). In this section we will not consider segregation, in order to maximize the size of the available samples.

From Table 59, the model that has the lowest AIC (= -294.7) and BIC (= -283.2) selects South, Central, NEast, Age70, Nonw70, Lgvpc70, Igr70, Lpcex70, Exhwy70, Ldebt70, Lpc70, and Lmedic70. All five models in Table 59 select South, Central, NEast, Age70, Nonw70, Lgvpc70, and Lpcex70. But the education variables are not selected in any model, which is not as expected. From Table 60, the model that with the lowest AIC (= -191.2) and BIC (= -185.5) selects Central, NEast, Lur70, Mfgs70, Lgvpc70, Pctax70, and Lpop70. All models in Table 60 select Central, NEast, Lur70, Lgvpc70, and Lpop70. Education and inequality variables are not selected by one model, but government variables (Lgvpc70, Pctax70 and Exedu70) are selected by various models..

From Table 61, the model that with the highest probability (= 0.0204) selects South, Central, NEast, Coll70, Age70, Nonw70, Lur70, Medsy70, Incla70, Igr70, Expo70, Ldebt70, and Lpop70. South (posterior t-probability = 0.096241), Central (posterior t-probability = 0.988209), NEast (posterior t-probability = 0.000), Coll70 (posterior t-probability = 0.000001), Age70 (posterior t-probability = 0.270972), Nonw70 (posterior t-probability = 0.210380), Lur70 (posterior t-probability = 0.233337), Medsy70 (posterior t-probability = 0.000037), Incla70 (posterior t-probability = 0.000213), Ldebt70 (posterior t-probability = 0.086741) and Lpop70 (posterior t-probability = 0.842216) are selected by all six models. This BMA approach, selects variables from every category and provides good models for us to analyze. The six models selected by BMA cover region (South, Central and Neast), education (Coll70, Medsy70), social

(Age70, Nonw70), inequality (Incla70), government (Lgvpc70, Igr70, Lpcex70, Expo70, and Ldebt70) and initial variables (Lpop70, Lur70, and Lpc70). These variables cover all categories and enable us to explain the relationship between each category and city growth (Table 62 is the MSA part).

For Table 62, the model that with the highest posterior probability (= 0.020) selects South (P.M.P. = 0.1349), Central (P.M.P. = 0.0001), NEast (P.M.P. = 0.0895), High70 (P.M.P. = 0.1563), Age70 (P.M.P. = 0.0000), Nonw70 (P.M.P. = 0.0000), Lgvpc70 (P.M.P. = 0.2448), Lpcex70 (P.M.P. = 0.0289), Ldebt70 (P.M.P. = 0.0835), Lpc70 (P.M.P. = 0.3279), and Edle70 (P.M.P. = 0.1588). The 20 models in Table 62 all select South, Central, NEast, Age70, Nonw70, Lpcex70, Ldebt70 and Edle70. Geweke's approach also selects from each category, as the BMA, but provide another good perspective for the analysis. It includes region, education, inequality, social, government and initial variables, which also provides a good model to analysis. From Table 63, the model that has the highest posterior probability (0.002) selects Central (P.M.P. = 0.2170), NEast (P.M.P. = 0.2889), Medsy70 (P.M.P. = 0.3041), Lur70 (P.M.P. = 0.5220), Incle70 (P.M.P. = 0.5740), Lpcex70 (P.M.P. = 0.2094), Exedu70 (P.M.P. = 0.6786), Nonw70 (P.M.P. = 0.5739) and Lpop70 (P.M.P. = 0.3231). For the MSA, Table 63 also covers most categories with inequality (P.M.P. for Incle70 and Incla70 are 0.5740 and 0.6908, respectively); probably this category is not the main concern for the MSA area, compared with other categories. For this MSA analysis, government expenses become important (they are important for the CITY) so the models select those variables frequently. The posterior marginal probability (P.M.P.) for these five Tables (Table 61-65) will attach in Appendix B for further reference (Table 66 and 67).

From Table 64, the model that with the highest posterior probability (0.0036) selects Central and Medsy70. No inequality and government variables are selected in Table 64. MBVS selects very few variables for each model, as can be seen from Table 57 and Table 58. Also this result does not cover every category, as the previous two Bayesian approaches did. As we see from the last section (government category), the observation that MBVS selects few variables when the explanatory variables are larger is confirmed in Table 64. Now, turning to Table 65, the model that has the highest posterior probability (0.0143) selects Central and NEast and no model contains more than four variables. From Table 65, many models contain one variable and the other models contain two or three variables. So MBVS becomes very conservative when the explanatory variables become larger with conservative prior settings.

The correlation coefficients of the P.M.P. among the three Bayesian procedures is in Table 68 (for CITY, all variables which come from Table 66) and Table 69 (for MSA, all variables which come from Table 67). From Table 68 and Table 69, the correlation between Geweke and MBVS is considered as high (= 0.52806 and 0.73422, respectively). This can be interpreted to mean that the relative rankings produced by the two procedures are considerably consistent. BMA is not highly correlated with these two procedures, which is not as I expect.

The Geweke procedure, with the priors used in this study, tend to select models with many more variables than the MBVS. This is, at least in part, due to the selected hyperparmeters. Decreasing δ in MBVS is likely to yield larger models. The high correlation among the P.M.P. of these two models suggests they may in fact yield similar results.

From an overall standpoint, the variables selected for the CITY and MSA data sets differ quite a bit. This may indicate that the underlying economic processes differ for the two. In the end, the Bayesian procedures are more flexible since they can systematically use expert knowledge that is available to the user. The largest disadvantage comes from computational complexity and possible uncertainty over the selected priors.

CHAPTER FIVE

CONCLUSION AND FUTURE EXTENSION

In this dissertation, I used several Bayesian variables and model selection methods to explore an issue in urban economics. Compared to those traditional methods (AIC or BIC), the Bayesian methodology is very flexible for researchers to implement. The biggest advantage for the Bayesian methodology is that it allows using prior information on those variables based on the researcher's expert knowledge. Using this unique prior knowledge, the variables after selection can provide more information about the relationship between variables. The usual caveat applies: The better prior information leads to better overall results in Bayesian analysis. In this final chapter, I will discuss several issues regarding the variable selection problem and possible future improvements in variable selection. The biggest disadvantage is the computational complexity of the procedures, which will surely discourage many from using any of the Bayesian procedures considered. This can be expected to improve as commercial software becomes available and as computational speeds increase.

5.1 Prior Setting for Bayesian Variable Selection

For these three Bayesian variable selection methods, I implemented easily specified prior information (mostly in uniform prior form) for the parameters because my purpose is to see how these methods work with the economic data. But these prior forms may not reflect the best available information about the effects of these variables. This prior setting remains one of the most difficult parts of using the Bayesian approach because these settings have large effects on the final selection of the model. Using subjective prior forms may be a solution because it assumes the researcher has exact knowledge about the variables and parameters. But this subjective prior form relies on the past experience of a researcher who has worked with those variables and parameters. So before choosing a Bayesian prior for the parameters of a model, one should avoid oneself of what ever expert knowledge there is. This expertise in prior settings is not expected to eliminate bias (which is probably impossible) but it should reduce the bias and result in more reliable model selection.

5.2 Posterior Distribution

The posterior distribution is another important part because we need to draw samples from the posterior distribution to compute the posterior probability. It is always difficult to ascertain the exact posterior distribution for the parameters that we are interested in. However, recent advance in computational Bayesian analysis has given us means to approximate these with reasonable accuracy.

5.3 MCMC Algorithm

In this dissertation, the MCMC algorithm I implement is Gibbs sampling because it is easy to implement and understand. But there are other, more efficient MCMC algorithm we should consider; e.g., the Metropolis-Hastings algorithm. The Metropolis-Hastings algorithm is more suitable than Gibbs sampling if the generating parameters do not have a standard distribution but we can assume they have a known kernel of density

(Monfort & Dijk, 1996). This algorithm should provide us another good alternative algorithm in this variable selection problem because George et al. (1997) already implemented it in the Bayesian variable selection problem.

If the data set is very large and the collinearity is high among the explanatory variables, the number of iterations needs to be larger in order to achieve the numerical accuracy (Geweke, 1994). Assured of convergence, the posterior estimates will be more consistent. So these two problems need be considered when we implement the Bayesian approaches to solve the variable selection problem.

5.4 Software

The software programs I used to carry out these three Bayesian approaches are Gauss (1995) and Matlab (1999). They are adapted from the website (except for Gauss, which is adapted from Adkins et al. 1999) and revised for this dissertation. The website for BMA is www.spatial-econometrics.com and for MBVS is stat.tamu.edu/~mvannucci/webpages/codes.html. They are capable of carrying out many Bayesian computations and are easy to use. Other software can also carry out these Bayesian computations, like BUGS (Bayesian Using Gibbs Sampling), O-Matrix or C++. BUGS and O-Matrix (the light version) are free and could be downloaded from the internet. C++ is a special (but powerful) computer language and, although more knowledge is necessary to use it, it should perform well also.

5.5 Future Extension and Summary

This dissertation tries to reveal the problem of variable and model selection in economics. Compare the results with GSS, the Bayesian methods are accounting more variables into the models which we expect. For these three Bayesian approaches, BMA and Geweke's approach seem to be more reasonable to the MBVS since the empirical results from Chapter 4 suggest. But with more flexible prior setting, MBVS should perform reasonable well as the other two approaches in the future.

The variable selection problem has been gaining more attention during the last 20 years and many methods have become available for us to implement. As information becomes easier to get, we will have larger data sets and many variables to analyze. So the variable selection problem will become more important, not only for the researchers but also for corporate decision-makers. As this trend grows, researchers will need to cope with many tasks in order to solve more advanced problems. For example, the posterior estimates ($E(\beta_i | M_j)$) for each parameter under different models and the posterior standard deviation estimates. When these two estimates are obtained, we can construct the posterior confidence inference and make inferences about the parameters. This will be a major task for the Bayesian variable selection approaches in the near future. Also new developments for the software to carry out the complex Bayesian computation is also desired since the existing software is not so efficient (because they take a long time to compute). Finally, we need to be more cautious about the danger (such as misleading decisions, inconsistencies etc.) that these Bayesian approaches will cause, as George (2000) states "Our enthusiasm for the development of promising new procedures must be carefully tempered with cautious warnings of their potential pitfalls".

REFERENCES

Adkins, L., Moomaw, R., & Tien, J. C. (1999). *Bayesian Model Selection: An Application in Urban Economics.* Paper presented at the meeting of the Southern Economic Association, New Orleans, LA.

Atkinson, R. (1996). The Rise of the Information-Age Metropolis. *The Futurist*, July-August, 41-46.

Amemiya, T. (1980). Selection of Regressors. *International Economic Review*, 21, 331-354.

Barro,R. J. (1998). *Determinants of Economic Growth*, Cambridge, MA: MIT Press.

Barro, R. J. & Sala-Martin, X. (1992). Convergence, *Journal of Political Economy,* 100(2), 223-251.

Bradbury, K.L., Downs, A. & Small, K. (1982). *Urban Decline and the Future of American Cities*, Washington, DC: Brookings Institution

Brown, P. J., Fearn, T., & Vannucci, M. (1998). Multivariate Bayesian Variable Selection and Prediction, *Journal of the Royal Statistical Society (Series B),* 60(3), 627-641.

Brown, P. J., Fearn, T., & Vannucci, M. (1999). The Choice of Variables in Multivariate Regression: A Non-conjugate Bayesian Decision Theory Approach, *Biometrika*, 86(3), 635-648.

Burnham, K. P. & Anderson, D. R. (1998). *Model Selection and Inference: A Practical Information-Theoretic Approach*, New York, NY: Springer-Verlag New York Inc.

Carlin, B. P. & Chib, S. (1993). Bayesian Model Choice via Markov Chain Monte Carlo, *Journal of the Royal Statistical Society (Series B*), 55, 473-484.

Chipman, H. (1996). Bayesian Variable Selection with Related Predictors, *The Canadian Journal of Statistics*, 24, 17-36.

Clyde, M., Desimone, H., & Parmigiani, G. (1996). Prediction via Orthogonalized Model Mixing, *Journal of the American Statistical Association*, 91, 1197-1208.

Clyde, M. & Parmigiani, G. (1998). Protein Construct Storage: Bayesian Variable Selection and Prediction with Mixtures, *Journal of Biopharmaceutical Statistics*, 8(3), 431-443.

Culter, D. M. & Glaeser, E. L. (1997). Are Ghettos Good or Bad? *The Quarterly Journal of Economics*, 112, 827-872.

Culter, D. M. & Vigdor, (1999). The Rise and Decline of the American Ghetto, *Journal of Political Economy*, 107, 455-506.

Dieblot, J. & Robert, C. P. (1994). Estimation of Finite Mixture Distributions through Bayesian Sampling, *Journal of the Royal Statistical Society (Series B)*, 56(2), 363-375.

Dornbusch, R. & Fischer, S. (1994). *Macroeconomics*, New York, NY: McGraw-Hill.

Gauss (1995). *Version 3.2.13*, Aptech Systems, Inc., Maple Valley, WA.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). *Bayesian Data Analysis*, New York, NY: Chapman & Hall.

George, E. I. (1995). *Bayesian Model Selection.* (CCS Tech. Rep.). Austin, TX: University of Texas-Austin.

George, E. I. (2000). The Variable Selection Problem, *Journal of the American Statistical Association*, 95(452), 1304-1308.

George, E. I. & Foster, D. P. (1997). *Calibration and Empirical Bayes Variable Selection.* (CSS Tech. Rep.). Austin, TX: University of Texas-Austin.

George, E. I. & McCulloch, R. E. (1993). Variable Selection Via Gibbs Sampling. *Journal of the American Statistical Association*, 88, 881-889.

George, E. I. & McCulloch, R. E. (1995). Stochastic Search Variable Selection. In W. R. Gillks, S. Richardson, & D. J. Spiegelhalter (Eds). *Markvo Chain Monte Carlo In Practice* (pp. 203-214), London: Chapman & Hall.

George, E. I., McCulloch, R. E., & Tsay, R. S. (1996). Two Approaches to Bayesian Model Selection with Application. In D. Berry, K. Chaloner, & J. Geweke (Eds). *Bayesian Analysis in Statistics and Econometrics : Essays in Honor of Arnold Zellner* (pp. 339-348), New York, NY: Wiley.

George, E. I. & McCulloch, R. E. (1997). *Approaches For Bayesian Variable Selection.* (CSS Tech. Rep.). Austin, TX: University of Texas-Austin.

Geweke, J. (1994). *Variable Selection and Model Comparison in Regression.* (Tech. Rep.). Minneapolis, MN: University of Minnesota, School of Statistics.

Geweke, J. (1999). Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication (with discussion and rejoinder). *Econometric Reviews*, 18 , 1-126.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in Cities. *Journal of Political Economy*, 100(6), 1126-1152.

Glaeser, E. L., Scheinkman, J. A., & Schleifer, A. (1995). Economic Growth in A Cross-Section of Cities. *Journal of Monetary Economics*, 36, 117-143.

Green, P. J. (1995). Reversible jump Markov Chain Monte Carlo Computation and Bayesian Model Determination, *Biometrika*, 82, 711-732.

Green, W. H. (1993). *Econometric Analysis* (2nd Ed.), Englewood Cliffs, NJ: Prentice-Hall.

Hoeting, J. A. (1994). *Accounting for Model Uncertainty in Linear Regression.* Unpublished doctor's dissertation, Department of Statistics, University of Washington, Seattle, Washington, USA.

Hoeting, J. A., Raftery, A. E., & Volinsky, C. T. (1998). *Bayesian Model Averaging, Tech Report 335*, Department of Statistics, University of Washington, Seattle, Washington, USA.

Jacobs, J. (1984). *Cities and Wealth of Nations*, New York, NY: Vintage Books.

Judge, G. G., Hill, C. R., Griffiths, W. E., Lutkepohl, H., & Lee, T. (1988). *Introduction to the Theory and Practice of Econometrics* (2nd ed.). New York, NY: Wiley.

Kempen, R. V. & Ozuekren, A. S. (1998). Ethic segregation in cities: New forms and explanations in a dynamic world, *Urban Economics*, 35(10), 1631-1656.

Klein, L. (1950). *Economic Fluctuations in the United States 1921-1941*, New York, NY: John Wiley & Sons, Inc.

Krugman, P. (1991). Increasing Returns and Economic Geography, *Journal of Political Economy*, 99(3), 483-499.

Kuo, L. & Mallick, B. (1994). *Variable Selection for Regression Models* (Tech. Rep. No. 9426). Storrs, CT: University of Connecticut, Department of Statistics.

Lucas, R. E. Jr. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics*, 22, 3-42.

Madigan, D. & York, J. (1995). Bayesian Graphical Models for Discrete Data, *International Statistical Review*, 63, 215-232.

Matlab (1999). *Version 5.3*, The Math Works, Inc., Natick, MA.

McDonald, J. F. (1997). *Fundamentals of Urban Economics*. Upper Saddle River, NJ: Prentice-Hall.

Mieszkowski, P. & Mills, E. S. (1993). The Cause of Metropolitan Suburbanization, *Journal of Economic Prespectives*, 7(3), 135-147.

Mills, E. S. & Hamilton, B. W. (1994). *Urban Economics* (5th Ed.). New York, NY: HarperCollins College Publisher.

Mills, E. S. & Lubuele, L. S. (1997). Inner Cities, *Journal of Economic Literature*, 35, 727-756.

Mitchell, T. J. & Beauchamp, J. J. (1988). Bayesian Variable Selection in Linear Regression. *Journal of American Statistical Association*, 83, 1023-1032.

Monfort, A. & Dijk, H. K. V. (1995). Simulation-Based Economics, In H. K. V. Dijk, A. Monfort, & B. W. Brown (Eds). *Econometric Inference Using Simulation Techniques* (pp. 1-20). New York, NY: John Wiley & Sons, Inc.

Moulton, B. R. (1991). A Bayesian Approach to Regression Selection and Estimation with Application to a Price Index for Radio Services, *Journal of Econometrics*, 49, 169-193.

Pearce, D. W. (1996). The MIT Dictionary of Modern Economics (4th Ed.). Cambridge, MA: The MIT Press.

Raftery, A. E. (1995). Bayesian Model Selection in Social Research. In P. V. Marsden (Eds.). *Sociological Methodology* (pp. 111-196). Cambridge, MA: Blackwells.

Raftery, A. E., Madigan, D., & Hoeting, J. A. (1997). Bayesian Model Averaging for Linear Regression Models. *Journal of American Statistical Association*, 92, 179-191.

Raftery, A. E., Madigan, D., Hoeting, J. A., & Volinsky C. T. (1998). *Bayesian Model Averaging*. (Tech. Rep. No. 335). Seattle, WA: University of Washington, Department of Statistics.

Rauch, J. E. (1993). Productivity Gains from Geographic Concentration of Human Capital : Evidence from the Cities. *Journal of Urban Economics*, 34, 380-400.

Richard, J. F. (1995). Bayesian Model Selection and Prediction with Empirical Applications: Discussion, *Journal of Econometrics*, 69, 337-349.

Sawa, T. (1978). Information Criteria for Discriminating Among Alternative Regression Models. *Econometrica*, 46, 1273-1291.

Schwarz, G. (1978). Estimating the Dimension of a Model, *The Annals of Statistics*, 6, 461-464.

Seber, G.A.F. (1984). *Multivariate Observations*, New York, NY: John Wiley & Sons.

Simon, C. J. (1996). The Talk of Town: Human Capital, Information and the Growth of English Cities, 1861-1961, *Exploration in Economic History*, 33, 384-413.

Simon, C. J. (1998). Human Capital and Metropolitan Employment Growth, *Journal of Urban Economics*, 43, 223-243.

Smith, M. & Kohn, R. (1996). Nonparametric Regression using Bayesian Variable Selection, *Journal of Econometrics*, 75, 317-343.

South, S. J. & Crowder, K. D. (1997). Residential Mobility between Cities and Suburbs: Race, Suburbanization, and Back-to-the City Moves, *Demography*, 34, 525-538.

Steinacker, A. (1998). Economic Restructuring of Cities, Suburbs and Non-metropolitan Areas. *Urban Affairs Review*, 34(2), 212-240.

Taeuber, A. & Taeuber, K. (1965). *Negroes in Cities,* New York, NY: McGraw-Hill.

Thisted, R (1988). *Elements of Statistical Computing-Numerical Computation*, New York, NY: Chapman and Hall.

Voith, R. (1998). Do Suburbs Needs Cities ? *Journal of Regional Science*, 38(3), 445-464.

Zellner,A. (1986). On Assessing Prior Distributions and Bayesian Regression Analysis With g Prior Distributions. In P.K. Goel & A. Zellner (Eds). *Bayesian Inference and Decision Techniques – Essays in Honor of Bruno de Finetti* (233-243), Amsterdam: North-Holland.

APPENDIX A

Appendix A

Description of Variables

The following variables are obtained from County and City Data Book (1972&1994), City of the U.S website (www.policy.rutgers.edu/cupr/sonc.htm), Government information share website (www.kerr.orst.edu) and MSA segregation index is from Cutler, Glaser and Vigdor segregation index website (www.pubpol.duke.edu/~jvigdor/segregation/index/htm). For the city file, it has 77 cities. For MSA file, it has 75 cities because two are combined as one MSA (Los Angeles-Long Beach and Minneplois-St. Paul). The variables are index by 70 as they are observations from 1970. The central city file will call as CITY and the MSA as described in Chapter Four.

For the CITY data set, it includes:

Lpop70: The log of city and MSA population for 1970 which from County and City Data Book.

Region: The region which is defined by Taeuber et al (1965), divided into four region: South, Central, NEast (North East) and West.

Education: This category has three variables: High70 (the percentage of population of 1970 whose age are over 25 with 12+ years of education), Coll70 (have 16+ years of education) and Medsy70 (median school years for population 25 years or over).

Age70: This is the percentage of population who is over 65 years old.

Mfgs70: The percentage of workers working in manufacturing industries.

Nonw70: The percentage of population, which is not white.

Lur70: log of unemployment rate.

Incle70 and Incla70: The percentage of population who has income less than $3000(Incle70) and more than $25000(Incla70).

Lgvpc70: Log of per-capital government revenue

Igr70: inter-government revenue

Lpcex70: Log of per-capital government expenditure

Exhwy70, Expo70 and Exss70: Percentage of government expenditure spends on highway, police protection and sanitation service, respectively.

Ldebt70: Log of debt that government is outstanding.

Lpc70: log of the per-capita money income.
Seg70: Segregation index for city and MSA, which is defined as Taeuber et al (1965).

Lpmed70: Medsy70 weighted by multiplying Lpop70.

Weseg70: Seg70 weighted by multiplying Nonw70.

Edle70: Percentage of population has 5 or less years in education

Lmedic70 : Log of per-capita median money income

For the MSA data set, it includes:

Lpop70: log of population of MSA of 1970

Region: Same as CITY, has South, Central, and NEast.

Education: Same as CITY, has High70, Coll70, and Medsy70 of MSA

Age70: Same as CITY, percentage of people 65 or older of MSA

Mfgs70: Same as CITY, percentage of workers working in manufacturing industries of MSA

Nonw70: Same as CITY, percentage of population which is not white of MSA

Lur70: Same as CITY, log of unemployment rate of MSA

Incle70 and Incla70: Defined the same as CITY, percentage of population who has income less than $3,000 and more than $25,000 of MSA

Lgvpc70: log of government total revenue of MSA

Igr70: inter-government revenue of MSA

Lpcex70: log of per-capita government expenditure of MSA

Exhwy70, Exedu70, and Exheal70: percentage of government expenditure spends on highway, education, and health care of MSA.

Pctax70: percentage of government revenue from tax of MSA

Seg70: segregation index for MSA, defined the same as CITY

Weseg70: Seg70 weighted by multiplying Lpop70

Lpc70: log of per-capita money income of MSA

The following are 14 cities do not have dissimilarity index available:
1.Anchorage, AK
2.Hartford, CT
3.Wilimington, DE
4.Boise, ID
5.Kansas City, KS
6.Portland, ME
7.Billings, MT
8.Manchester, NH
9.Fargo, ND
10.Columbia, SC
11.Sioux Falls, SD
12.Burlington, VT
13.Charleston, WA
14.Cheyenne, WY

The following are 12 MSAs do not have dissimilarity index available:
1.Anchorage, AK
2.Santa Ana, CA
3.Boise, ID
4.Kansas City, KS
5.Portland, ME
6.Billings, MT
7.Manchester, NH
8.Fargo, ND
9.Sioux Falls, SD
10.Burlington, VT
11.Virginia Beach, VA
12.Cheyenne, WY

List of variables over which the selection exercises are conducted. The GSS (1995) table number and generic name for the variable set is given for each.

1.Table 4 (manufacturing):
Lpop70, Lpc70, Mfgs70, South, Central and NEast.

2.Table 5 (unemployment):
Lpop70, Lpc70, Mfgs70, Lur70, South, Central and NEast.

3.Table 6 (education):
Lpop70, Lpc70, Mfgs70, Lur70, High70, Coll70, Medsy70, Lpmed70, South, Central, and NEast.

4.Table 8 (inequality):
Lpop70, Lpc70, Mfgs70, Lur70, High70, Medsy70, Incle70, Incla70, Edle70, Lmedic70,

South, Central and NEast.

5. Table 9 (social characteristics):
Lpop70, Lpc70, Age70, Nonw70, Lur70, Mfgs70, Medsy70, Seg70, Weseg70, South, Central, and NEast.

6. Table 10 (Government expenditure and revenue):
Lpop70, Lpc70, Medsy70, Lur70, Mfgs70, Lgvpc70, Igr70, Lpcex70, Exhwy70, Expo70, Exss70, Exedu70, Exheal70, Ldebt70, Pctax70, South, Central, and NEast.

APPENDIX B

Appendix B: Table for the Analysis

Table 1
Mean and Standard Deviation for CITY Variables

Variables	Mean	Std. dev.	Minimum	Maximum
Population growth	0.115	0.341	-0.449	1.549
South	0.312	0.466	0.000	1.000
Central	0.259	0.441	0.000	1.000
NEast	0.182	0.388	0.000	1.000
High70	54.575	10.201	33.100	76.100
Coll70	12.047	3.643	4.400	20.700
Age70	10.334	2.675	2.000	15.000
Mfgs 70	19.339	8..598	3.400	37.500
Nonw70	20.356	16.054	0.300	72.100
Lur70	1.522	0.277	0.956	2.116
Medsy70	11.878	0.753	9.600	12.800
Incle70	10.117	2.872	1.600	19.000
Incla70	4.678	2.438	1.700	18.000
Lgvpc70	5.046	0.857	1.872	6.732
Igr70	20.864	11.732	3.200	53.900
Lpcex70	5.020	0.605	4.111	6.746
Exhway70	10.863	5.841	1.100	28.800
Expo70	21.547	7.328	5.700	41.200
Exss70	11.358	6.130	2.300	26.800
Ldebt70	4.658	1.358	1.569	9.070
Lpc70	8.097	0.130	7.792	8.508
Edle70	5.290	2.731	1.400	15.300
Lmedic70	9.146	0.120	8.782	9.500
Lpop70	1.743	0.178	1.243	2.194

Table 1A
Proportions of Coefficient Variance Associated with Each Characteristic Root
(For CITY)

Eigenvalue	CONDITIONAL INDEX	INTERCEP	SOUTH	CENTRAL	NEAST	HIGH70	COLL70
19.67866	1.00000	0.0000	0.0001	0.0001	0.0001	0.0000	0.0000
1.23483	3.99203	0.0000	0.0092	0.0059	0.0742	0.0000	0.0001
1.13155	4.17023	0.0000	0.0330	0.0707	0.0003	0.0000	0.0000
0.52917	6.09818	0.0000	0.0092	0.0605	0.0045	0.0001	0.0028
0.33710	7.64049	0.0000	0.0224	0.0045	0.0603	0.0000	0.0006
0.22965	9.25683	0.0000	0.0573	0.0081	0.0067	0.0000	0.0002
0.19014	10.17325	0.0000	0.0758	0.1519	0.1467	0.0000	0.0028
0.14175	11.78267	0.0000	0.0094	0.0037	0.0001	0.0001	0.0032
0.12925	12.33898	0.0000	0.0069	0.0000	0.0254	0.0000	0.0064
0.10525	13.67356	0.0000	0.0001	0.0225	0.0088	0.0001	0.0066
0.07828	15.85521	0.0000	0.0168	0.0525	0.0057	0.0000	0.0012
0.06479	17.42812	0.0000	0.0041	0.0063	0.0097	0.0001	0.0145
0.04755	20.34258	0.0000	0.0406	0.0546	0.0582	0.0000	0.0062
0.03957	22.29987	0.0000	0.0001	0.0370	0.0050	0.0002	0.0512
0.02181	30.03833	0.0000	0.0053	0.0018	0.0069	0.0001	0.3631
0.01495	36.28292	0.0001	0.0138	0.0200	0.0566	0.0045	0.0947
0.01100	42.29431	0.0000	0.0014	0.0416	0.0751	0.0031	0.0144
0.00875	47.43519	0.0000	0.4435	0.1937	0.1057	0.0003	0.0001
0.00298	81.20746	0.0000	0.0001	0.0062	0.0127	0.0308	0.1059
0.00176	105.63411	0.0001	0.0003	0.0017	0.0791	0.1406	0.0669
0.00084	152.31981	0.0030	0.1179	0.1171	0.0302	0.3087	0.0889
0.00030	254.49262	0.0043	0.0094	0.0000	0.0025	0.3045	0.0002
0.00003	746.83947	0.1317	0.0187	0.0824	0.1023	0.1649	0.1523
8.59239E-6	1513	0.8608	0.1046	0.0572	0.1235	0.0418	0.0176

AGE70	MFGS70	NONW70	LUR70	MEDSY70	INCLE70	INCLA70	LGVPC70	IGR70
0.0001	0.0001	0.0002	0.0000	0.0000	0.0000	0.0001	0.0000	0.0003
0.0001	0.0004	0.0020	0.0000	0.0000	0.0000	0.0002	0.0000	0.0043
0.0001	0.0005	0.0022	0.0000	0.0000	0.0001	0.0005	0.0000	0.0007
0.0000	0.0036	0.0428	0.0002	0.0000	0.0003	0.0051	0.0000	0.0001
0.0008	0.0070	0.0685	0.0000	0.0000	0.0004	0.0541	0.0000	0.0276
0.0000	0.0000	0.0522	0.0001	0.0000	0.0002	0.0096	0.0006	0.2306
0.0005	0.0067	0.0070	0.0043	0.0000	0.0003	0.0476	0.0001	0.0435
0.0001	0.0059	0.0401	0.0010	0.0000	0.0002	0.0103	0.0019	0.1232
0.0036	0.0435	0.0612	0.0015	0.0000	0.0001	0.0704	0.0001	0.0906
0.0128	0.0706	0.0170	0.0022	0.0000	0.0186	0.0707	0.0002	0.0259
0.0049	0.0859	0.1152	0.0000	0.0000	0.0005	0.0295	0.0000	0.0486
0.0384	0.0006	0.0208	0.0000	0.0000	0.0116	0.2228	0.0019	0.1614
0.0320	0.0690	0.0082	0.0036	0.0000	0.0050	0.0001	0.0012	0.0282
0.2374	0.1255	0.0002	0.0118	0.0000	0.0011	0.0043	0.0000	0.0086
0.1100	0.0544	0.0095	0.1001	0.0000	0.0095	0.0001	0.0307	0.0273
0.0026	0.0017	0.0310	0.2516	0.0017	0.0690	0.0027	0.0003	0.0212
0.0608	0.0056	0.0095	0.0001	0.0006	0.0805	0.0100	0.5060	0.0011
0.1364	0.0417	0.1368	0.5250	0.0000	0.4271	0.0135	0.0884	0.0016
0.0000	0.0093	0.0506	0.0155	0.0039	0.0044	0.0000	0.0139	0.0019

AGE70	MFGS70	NONW70	LUR70	MEDSY70	INCLE70	INCLA70	LGVPC	IGR70
0.0627	0.0407	0.2538	0.0142	0.0020	0.0054	0.0281	0.3086	0.0523
0.0025	0.2660	0.0099	0.0002	0.0028	0.0172	0.0117	0.0268	0.0224
0.0009	0.0029	0.0046	0.0465	0.8760	0.0049	0.1570	0.0027	0.0014
0.0492	0.0244	0.0017	0.0068	0.1052	0.0310	0.2048	0.0054	0.0011
0.2444	0.1340	0.0549	0.0152	0.0077	0.3126	0.0468	0.0111	0.0762

LPCEX70	EXHWAY70	EXPO70	EXSS70	LDEBT70	LPC70	LPOP70	EDLE70	LMEDIC70
0.0000	0.0002	0.0001	0.0002	0.0000	0.0000	0.0000	0.0001	0.0000
0.0000	0.0035	0.0004	0.0064	0.0000	0.0000	0.0000	0.0001	0.0000
0.0000	0.0015	0.0000	0.0001	0.0000	0.0000	0.0000	0.0016	0.0000
0.0000	0.0178	0.0003	0.0005	0.0003	0.0000	0.0000	0.0062	0.0000
0.0000	0.0029	0.0020	0.0187	0.0002	0.0000	0.0000	0.0010	0.0000
0.0000	0.0145	0.0185	0.0204	0.0005	0.0000	0.0000	0.0048	0.0000
0.0000	0.0191	0.0132	0.0032	0.0001	0.0000	0.0001	0.0080	0.0000
0.0001	0.3035	0.0067	0.0611	0.0034	0.0000	0.0004	0.0000	0.0000
0.0000	0.0004	0.0003	0.3375	0.0002	0.0000	0.0002	0.0046	0.0000
0.0001	0.0202	0.0015	0.1639	0.0051	0.0000	0.0009	0.0002	0.0000
0.0000	0.0230	0.0051	0.0004	0.0002	0.0000	0.0001	0.2582	0.0000
0.0001	0.0967	0.0217	0.1163	0.0167	0.0000	0.0037	0.0180	0.0000
0.0001	0.0525	0.4328	0.0711	0.0059	0.0000	0.0022	0.0091	0.0000
0.0000	0.1007	0.0051	0.0068	0.0305	0.0000	0.0041	0.0076	0.0000
0.0015	0.0555	0.0561	0.0455	0.0019	0.0000	0.0002	0.0836	0.0000
0.0072	0.1136	0.0003	0.0015	0.0509	0.0002	0.0005	0.0622	0.0001
0.0003	0.0257	0.0357	0.0088	0.0151	0.0000	0.0349	0.0158	0.0000
0.0000	0.0524	0.0025	0.0027	0.0386	0.0000	0.0002	0.0433	0.0000
0.0502	0.0012	0.0060	0.0517	0.5577	0.0000	0.6255	0.0258	0.0000
0.3494	0.0507	0.0767	0.0280	0.2062	0.0006	0.2626	0.0884	0.0001
0.5022	0.0162	0.3084	0.0144	0.0270	0.0080	0.0200	0.1265	0.0020
0.0019	0.0063	0.0048	0.0005	0.0158	0.0015	0.0004	0.0010	0.0031
0.0145	0.0000	0.0007	0.0052	0.0234	0.8833	0.0440	0.1980	0.0260
0.0723	0.0218	0.0011	0.0350	0.0003	0.1064	0.0000	0.0359	0.9688

Table 2
Mean and Standard Deviation for the MSA Variables

Variables	Mean	Std. dev.	Minimum	Maximum
Population growth	0.406	0.333	-0.126	1.612
South	0.320	0.470	0.000	1.000
Central	0.240	0.430	0.000	1.000
NEast	0.173	0.381	0.000	1.000
Medsy70	12.062	0.374	10.780	12.600
High70	57.420	7.309	40.090	75.900
Coll70	12.098	2.953	7.620	23.400
Lur70	1.425	0.279	0.963	2.110
Mgfs70	21.707	9.496	3.100	42.680
Incle70	8.821	2.819	4.400	16.340
Incla70	4.788	1.893	1.400	11.700
Lgvpc70	5.190	1.183	2.741	7.999
Igr70	34.532	13.993	13.390	119.010
Pctax70	49.958	10.913	14.400	76.090
Lpcex70	6.545	0.747	4.304	8.115
Exedu70	51.144	7.498	34.700	81.520
Exhway70	7.823	3.711	1.600	23.800
Exheal70	5.006	4.498	0.100	34.030
Lpc70	13.437	1.132	10.940	16.264
Age70	8.707	2.356	1.400	20.300
Nonw70	12.573	10.041	0.300	58.630
Lpop70	13.437	1.132	10.940	16.264

Table 2A
Proportions of Coefficient Variance Associated with Each Characteristic Root (For MSA)

Eigenvalue	CONDITIONAL INDEX	INTERCEP	LPOP70	SOUTH	CENTRAL	NEAST	MEDSY70
17.97163	1.00000	0.0000	0.0000	0.0001	0.0002	0.0001	0.0000
1.18978	3.88652	0.0000	0.0000	0.0525	0.0402	0.0302	0.0000
1.02318	4.19101	0.0000	0.0000	0.0000	0.1010	0.1254	0.0000
0.44361	6.36492	0.0000	0.0000	0.0273	0.0036	0.0013	0.0000
0.37467	6.92576	0.0000	0.0000	0.0116	0.1237	0.0491	0.0000
0.30193	7.71502	0.0000	0.0000	0.0726	0.0173	0.0424	0.0000
0.18018	9.98705	0.0000	0.0000	0.0868	0.0625	0.0135	0.0000
0.12987	11.76366	0.0000	0.0000	0.0521	0.0046	0.0001	0.0000
0.12112	12.18117	0.0000	0.0000	0.0296	0.0072	0.0001	0.0000
0.10744	12.93335	0.0000	0.0000	0.0137	0.1975	0.2109	0.0000
0.05526	18.03417	0.0000	0.0000	0.0037	0.0041	0.0065	0.0000
0.02966	24.61424	0.0000	0.0000	0.0330	0.0865	0.1745	0.0000
0.02093	29.30629	0.0000	0.0012	0.0046	0.0000	0.0309	0.0000
0.01652	32.98338	0.0000	0.0002	0.0497	0.0431	0.0017	0.0000
0.01337	36.66013	0.0000	0.0000	0.0019	0.0001	0.0082	0.0000
0.00956	43.35576	0.0001	0.0004	0.4333	0.2239	0.1602	0.0000
0.00457	62.70277	0.0009	0.0035	0.0032	0.0002	0.0059	0.0014
0.00401	66.95691	0.0003	0.0193	0.0434	0.0138	0.0005	0.0021
0.00183	99.22846	0.0066	0.0831	0.0538	0.0043	0.0595	0.0013
0.00049	190.75306	0.0134	0.8791	0.0000	0.0348	0.0018	0.0089
0.00033	232.35226	0.0609	0.0119	0.0266	0.0142	0.0507	0.1175
0.00006	526.71878	0.9178	0.0012	0.0002	0.0173	0.0262	0.8686

HIGH70	COLL70	LUR70	MFGS70	INCLE70	INCLA70	LGVPC70	IGR70	PCTAX70
0.0000	0.0001	0.0000	0.0001	0.0000	0.0001	0.0000	0.0002	0.0000
0.0000	0.0000	0.0000	0.0005	0.0005	0.0003	0.0000	0.0002	0.0002
0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0004	0.0000
0.0000	0.0000	0.0001	0.0002	0.0006	0.0010	0.0000	0.0058	0.0006
0.0002	0.0028	0.0019	0.0102	0.0002	0.0009	0.0000	0.0002	0.0005
0.0000	0.0016	0.0003	0.0065	0.0002	0.0110	0.0001	0.0001	0.0000
0.0000	0.0022	0.0001	0.0002	0.0002	0.0182	0.0006	0.0050	0.0001
0.0000	0.0023	0.0021	0.0009	0.0187	0.0444	0.0001	0.0720	0.0029
0.0000	0.0056	0.0017	0.0054	0.0062	0.0278	0.0002	0.4017	0.0057
0.0001	0.0083	0.0040	0.2056	0.0070	0.0000	0.0022	0.0036	0.0003
0.0007	0.0160	0.0001	0.0609	0.0040	0.0641	0.0081	0.0165	0.0053
0.0000	0.0883	0.2182	0.0039	0.0169	0.0319	0.0000	0.1832	0.0357
0.0001	0.0488	0.0152	0.0209	0.0005	0.2763	0.0550	0.0525	0.0419
0.0008	0.2860	0.0001	0.3030	0.2301	0.0136	0.0018	0.0028	0.1155
0.0032	0.0122	0.0026	0.0088	0.1164	0.0039	0.0041	0.1251	0.5803
0.0006	0.2256	0.4473	0.0184	0.2354	0.3944	0.0001	0.1119	0.0464
0.0373	0.0263	0.0686	0.1997	0.0152	0.0007	0.1067	0.0118	0.0182
0.0638	0.1912	0.1525	0.0009	0.0158	0.0781	0.0399	0.0005	0.0130
0.2755	0.0346	0.0006	0.0594	0.1388	0.0057	0.1224	0.0025	0.0000
0.0810	0.0005	0.0239	0.0620	0.0816	0.0000	0.6090	0.0028	0.0046
0.1097	0.0013	0.0176	0.0324	0.0940	0.0023	0.0031	0.0002	0.0969
0.4271	0.0465	0.0430	0.0002	0.0176	0.0251	0.0467	0.0007	0.0317

LPCEX70	EXEDU70	EXHWY70	EXHEAL70	LPC70	AGE70	NONW70
0.0000	0.0000	0.0002	0.0005	0.0000	0.0001	0.0003
0.0000	0.0000	0.0004	0.0116	0.0000	0.0000	0.0071
0.0000	0.0000	0.0002	0.0003	0.0000	0.0000	0.0011
0.0000	0.0000	0.0166	0.3941	0.0000	0.0000	0.0231
0.0000	0.0001	0.0014	0.0548	0.0000	0.0000	0.1403
0.0000	0.0002	0.0112	0.0157	0.0000	0.0038	0.2683
0.0001	0.0001	0.2796	0.0958	0.0000	0.0002	0.0374
0.0000	0.0000	0.1461	0.0011	0.0000	0.0429	0.0177
0.0000	0.0009	0.0029	0.0884	0.0000	0.0006	0.0003
0.0001	0.0004	0.0075	0.0648	0.0000	0.0009	0.0046
0.0000	0.0351	0.0705	0.0032	0.0000	0.1201	0.0104
0.0001	0.0000	0.0023	0.0073	0.0000	0.1434	0.0072
0.0219	0.0086	0.0216	0.0046	0.0001	0.1772	0.0182
0.0287	0.0132	0.0052	0.0015	0.0002	0.0004	0.0287
0.0037	0.0434	0.0066	0.1499	0.0004	0.3705	0.0746
0.0126	0.1148	0.0041	0.0031	0.0004	0.0449	0.1385
0.1045	0.6193	0.2514	0.0140	0.0045	0.0265	0.0033
0.4034	0.0126	0.0866	0.0003	0.0015	0.0215	0.0000
0.0196	0.0593	0.0778	0.0200	0.0437	0.0003	0.0511
0.2774	0.0484	0.0002	0.0506	0.1693	0.0042	0.1373
0.0104	0.0069	0.0045	0.0182	0.6739	0.0213	0.0073
0.1174	0.0368	0.0031	0.0001	0.1060	0.0212	0.0231

Table 3
CITY Population Growth and Manufacturing

Variables included	R^2	*Adjusted* R^2	*AIC*	*BIC*	*SBC*
Lpop70,Mfgs70,South, Central and NEast	0.489	0.453	-206.328	-203.203	-192.265
Mfgs70,South,Central, NEast	0.472	0.442	-205.811	-203.218	-194.092
Lpop70,LPc70,Central, NEast ,Mfgs70	0.465	0.435	-203.913	-200.344	-188.850
Lpop70,Mfgs70,South, Central ,NEast,Lpc70	0.491	0.450	-204.662	-201.282	-188.255
LPc70,Mfgs70,South, Central, NEast	0.472	0.435	-203.811	-201.097	-189.748

Table 3A
Proportion of Coefficient Variance Associated with Each Characteristic Root
(For CITY, CITY growth and manufacturing)

Eigenvalue	CONDITIONAL INDEX	INTERCEP	SOUTH	CENTRAL	NEAST	MFGS70	LPC70	LPOP70
4.67533	1.00000	0.0000	0.0050	0.0056	0.0045	0.0043	0.0000	0.0010
1.01731	2.14378	0.0000	0.2090	0.0699	0.0817	0.0014	0.0000	0.0000
1.00007	2.16218	0.0000	0.0007	0.1724	0.2485	0.0000	0.0000	0.0000
0.19536	4.89208	0.0001	0.5613	0.5265	0.4131	0.0006	0.0001	0.0090
0.09453	7.03274	0.0001	0.0071	0.1730	0.1531	0.8424	0.0001	0.0024
0.01732	16.42964	0.0015	0.0037	0.0069	0.0173	0.0807	0.0013	0.8936
0.00008	231.56760	0.9983	0.2132	0.0458	0.0819	0.0706	0.9985	0.0941

Table 4
MSA Population Growth and Manufacturing

Variables included	R^2	*Adjusted* R^2	*AIC*	*BIC*	*SBC*
Lpop70,Central,NEast	0.283	0.253	-182.767	-180.269	-173.497
Lpop70,Mfgs70,Central, NEast	0.300	0.260	-182.541	-179.670	-170.954
Lpop70,Mfgs70,Central	0.274	0.243	-181.835	-179.456	-172.565
Lpop70,Mgfs70,Lpc70, Central	0.292	0.251	-181.698	-178.966	-170.110
Lpop70,Lpc70,Mfgs70, Central,NEast	0.306	0.255	-181.278	-178.132	-167.375

Table 4A
Proportion of Coefficient Variance Associated with Each Characteristic Root
(For MSA, MSA growth and manufacturing)

Eigenvalue	CONDITIONAL INDEX	INTERCEP	SOUTH	CENTRAL	NEAST	MFGS70	LPC70	LPOP70
4.68399	1.00000	0.0000	0.0061	0.0055	0.0044	0.0042	0.0000	0.0002
1.01350	2.14980	0.0000	0.2172	0.0287	0.1618	0.0011	0.0000	0.0000
1.00003	2.16422	0.0000	0.0205	0.2500	0.1573	0.0000	0.0000	0.0000
0.21234	4.69671	0.0003	0.6112	0.4543	0.3295	0.0083	0.0003	0.0019
0.08614	7.37416	0.0003	0.0818	0.2563	0.2992	0.8999	0.0003	0.0011
0.00368	35.65559	0.0292	0.0018	0.0001	0.0122	0.0865	0.0221	0.9842
0.00032	121.22002	0.9702	0.0613	0.0051	0.0355	0.0000	0.9773	0.0125

Table 5
CITY Population Growth and Manufacturing (BMA)

Variables included	Model posterior probability
South,NEast,Lpc70,Lpop70	0.233
South,Mfgs70,Lpc70,Lpop70	0.213
South,Lpc70	0.062
South,Lpop70	0.056
South,NEast	0.047
South,Mfgs70	0.046
South,Central	0.045

Dependent variable: city population growth form 1970 to 1990
$v = 4.000, \lambda = 0.250$ and $\phi = 3.000$

Posterior t-probability: South=0.0003, Central=0.8833, NEast=0.5023, Mfgs70=0.4887,Lpc70=0.0092, Lpop70=0.0113

Table 6
MSA population growth and manufacturing (BMA)

Variables	Model posterior probability
South,NEast	0.102
NEast,Lpc70,Lpop70	0.090
South,Lpc70	0.068
South,Mfgs70	0.062
NEast,Lpc70	0.054
South,Lpop70	0.050

Posterior t-probability: South=0.0809, Central=0.8456, NEast=0.1452, Mfgs70=0.7493,Lpc70=0.1570, Lpop70=0.4993

Table 7
CITY Population Growth and Manufacturing (Geweke's)

Variables	Model posterior probability
South,Central,NEast,Mfgs70,Lpc70,Lpop70	0.495
South,Central,NEast,Mfgs70,Lpc70	0.319
Central,NEast,Lpc70,Lpop70	0.094
Central,NEast,Mfgs70,Lpc70	0.090

P.M.P. of variables: South=0.1851, Central=0.0001, NEast=0.0000, Mfgs70=0.0007, Lpc70=0.0000, Lpop70=0.4096

Table 8
MSA Population Growth and Manufacturing (Geweke's)

Variables	Model posterior probability
Central,NEast,Mfgs70,Lpc70	0.238
Central,NEast,Mfgs70,Lpc70	0.182
Central,Lpc70,Lpop70	0.164

P.M.P of variables: South=0.7256, Central=0.0206, NEast=0.1846, Mfgs70=0.2605, Lpc70=0.0000 Lpop70=0.3711

Table 9
CITY Population Growth and Manufacturing (MBVS)

Variables	Model posterior probability
Central,NEast,Mfgs70	0.109
Central,NEast,Mfgs70,Lpc70	0.109
South,Central,NEast,Mfgs70	0.066
South,Central,NEast,Mfgs70,Lpc70	0.065
Central,Mfgs70	0.054
Central,Mfgs70,Lpc70	0.054

P.M.P. of variables: South=0.6479, Central=0.3238, NEast=0.3450, Mfgs70=0.0810, Lpc70=0.5006, Lpop70=0.8353

Tables 10
MSA Population Growth and Manufacturing (MBVS)

Variables	Model posterior probability
Central,NEast	0.069
Central,NEast,Lpop70	0.068
Central,NEast,Lpc70,Lpop70	0.068
Central,NEast,Lpc70	0.067
Central,Lpop70	0.060
Central,Lpc70,Lpop70	0.060

P.M.P. of variables: South=0.5810, Central=0.2392, NEast=0.4520, Mfgs70=0.9881, Lpc70=0.5024, Lpop70=0.4374

Table 11
CITY Population Growth and Unemployment

Variables included	R^2	*Adjusted* R^2	*AIC*	*BIC*	*SBC*
Lur70,Mfgs70,South Central,NEast	0.475	0.438	-204.523	-201.575	-190.456
Lpop70,South,Central, NEast,Lur70,Mfgs70	0.492	0.447	-204.836	-201.279	-188.429
Mfgs70,South,Central, NEast	0.472	0.442	-205.811	-203.119	-194.092
Lpop70,Mfgs70,South, Central,NEast	0.489	0.453	-206.328	-203.082	-192.265
Lpop70,LPc70,Mfgs70 South,Central,NEast	0.491	0.450	-204.662	-201.139	-188.255

Table 12
MSA Population Growth and Unemployment

Variables included	R^2	*Adjusted* R^2	*AIC*	*BIC*	*SBC*
Lpop70,Central,Neast	0.283	0.253	-182.767	-180.179	-173.497
Lpop70,Mfgs70,Central, NEast	0.300	0.260	-182.541	-179.685	-170.954
Lpop70,Mfgs70,Central	0.274	0.243	-181.835	-179.451	-172.565
Lpop70,Lur70,Central, NEast	0.293	0.252	-181.779	-178.029	-170.191
Lpop70,Lpc70,Central, NEast	0.292	0.251	-181.698	-178.959	-170.110

Table 13
CITY Population Growth and Unemployment (BMA)

Variables included	Model posterior probability
South,Mfgs70,Lpc70,	0.151
South,Mfgs70,Lur70,Lpc70,Lpop70	0.139
South,Central,Mfgs70,Lpc70,Lpop70	0.081
South,NEast,Mfgs70,Lpc70,Lpop70	0.072
South,Central,NEast,Mfgs70,Lpc70,Lpop70	0.060
South,Lpc70	0.040

Posterior t-probability: South=0.0004, Central=0.7778, NEast=0.7085, Mfgs70=0.3138, Lur70=0.7070, Lpc70=0.0155, Lpop70=0.0208

Table 14
MSA Population Growth and Unemployment (BMA)

Variables	Model posterior probability
South,Lur70,Lpc70.Lpop70	0.064
Central,Lpc70,Lpop70	0.062
Central,Lur70,	0.050
South,Central,Mfgs70,Lpc70,Lpop70	0.047
South,Central,Lur70	0.047

Posterior t-probability: South=0.7487, Central=0.0924, NEast=0.8019, Lur70=0.1668, Mfgs70=0.8475, Lpc70=0.0931, Lpop70=0.3225

Table 15
CITY Population Growth and Unemployment (Geweke's)

Variables	Model posterior probability
South,Central,NEast,Mfgs70,Lpc70,Lpop70	0.328
South,Central,NEast,Mfgs70,Lpc70	0.231
South,Central,NEast,Mfgs70,Lur70,Lpc70,Lpop70	0.148
South,Central,NEast,Mfgs70,Lur70,Lpc70	0.099

P.M.P. of variables : South=0.1905, Central=0.0001, NEast=0.0002, Mfgs70=0.0020, Lur70=0.6919, Lpc70=0.0000, Lpop70=0.4084

Table 16
MSA Population Growth and Unemployment (Geweke's)

Variables	Model posterior probability
Central,NEast,Mfgs70,Lpc70	0.140
Central,NEast,Mfgs70,Lpc70,Lpop70	0.130
Central,NEast,Lur70,Mfgs70,Lpc70	0.097
Central,NEast,Lpc70,Lpop70	0.090

P.M.P. of variables: South=0.7310, Central=0.0228, NEast=0.1779, Lur70=0.6055, Mfgs70=0.2221, Lpc70=0.0000, Lpop70=0.4540

Table 17
CITY Population Growth and Unemployment (MBVS)

Variables	Model posterior probability
Central, NEast,Lpc70	0.071
Central,NEast	0.071
Central,NEast,Lur70,Lpc70	0.060
Central,NEast,Lur70	0.060
South,Central,NEast,Lpc70	0.040

P.M.P. of variables: South=0.6418, Central=0.1536, NEast=0.1854, Mfgs70=0.7857, Lur70=0.5361, Lpc70=0.4982, Lpop70=0.6902

Table 18
MSA Population Growth and Unemployment (MBVS)

Variables	Model posterior probability
Central, NEast	0.039
Central,NEast,Lpc70	0.038
Central,NEast,Lur70	0.033
Central,NEast,Lur70,Lpc70	0.033
Central,NEast,Lpop70	0.032
Central,NEast,Lpc70,Lpop70	0.032

P.M.P. of variables: South=0.5820, Central=0.2588, NEast=0.4581, Lur70=0.5385, Mfgs70=0.9895, Lpc70=0.5028, Lpop70=0.4808

Table 19
CITY Population Growth and Education

Variables included	R^2	*Adjusted* R^2	*AIC*	*BIC*	*SBC*
High70,Central,NEast	0.536	0.517	-217.800	-214.700	-208.400
Lmedsy70,High70, Central,NEast	0.540	0.514	-216.400	-213.100	-204.700
Lpop70,High70, Central,NEast	0.540	0.514	-216.400	-213.000	-204.700
LPc70,High70, Central,NEast	0.542	0.514	-216.900	-213.500	-205.200
High70,Coll70, Central,NEast	0.539	0.513	-216.300	-212.900	-204.500

Table 20
MSA Population Growth and Education

Variables included	R^2	*Adjusted* R^2	*AIC*	*BIC*	*SBC*
Mfgs70,Lpmed70, Central,NEast	0.305	0.266	-183.200	-180.100	-171.600
Lpc70,Mfgs70,Lpmed70, Central,NEast	0.325	0.277	-183.300	-179.800	-169.400
Lpop70,Central,NEast	0.283	0.253	-182.800	-180.300	-173.500
Lpc70,Lpmed70,Central, NEast	0.301	0.261	-182.700	-179.800	-171.100
Mfgs70,Lpmed70, Central	0.282	0.252	-182.700	-180.200	-173.400

Table 21
CITY population growth and education (BMA)

Variables included	Model posterior probability
South,NEast,Coll70,Mfgs70,Lpop70,Lpmed70	0.033
South,Central,NEast,Coll70,Mfgs70,Lpop70,Lpmed70	0.032
South,Coll70,Mfgs70,Lpop70,Lpmed70	0.019
South,Central,Coll70,Mfgs70,Lpop70,Lpmed70	0.016
South,Central,NEast,Coll70,Lur70,Lpop70,Lpmed70	0.016
South,Central,NEast,Lpop70,Lpmed70	0.015

Posterior t-probability: South=0.0024, Central=0.8333, NEast=0.4884, High70=0.7904,Coll70=0.1903, Mfgs70=0.4755, Lur70=0.7520, Medsy70=0.9751, Lpc70=0.7329, Lpop70=0.0001, Lpmed70=0.1484

Table 22
MSA population growth and education (BMA)

Variables	Model posterior probability
High70,Lur70,Mfgs70,Lpmed70	0.135
NEast,High70,Lur70,Mfgs70,Lpmed70	0.048
South,High70,Lur70,Mfgs70	0.024
South,High70,Mfgs70	0.015
High70,Mfgs70	0.015

Posterior t-probability: South=0.8047, Central=0.8309, NEast=0.9893, Medsy70=0.7973, Hgih70=0.0428, Coll70=0.8354, Lur70=0.4072, Mfgs70=0.2336, Lpc70=0.9569, Lpmed70=0.1039, Lpop70=0.9352

Table 23
CITY Population and Education (Geweke's)

Variables	Model posterior probability
Central,NEast,High70,Lpc70,Lpop70	0.100
Central,NEast,High70,Medsy70,Lpc70	0.073
Central,NEast,High70,Medsy70,Lpc70,Lpop70	0.069
Central,NEast,High70,Lpc70	0.059
Central,NEast,High70,Lur70,Lpc70,Lpop70	0.053
Central,NEast,High70,Lpc70,Lpop70	0.049
Central,NEast,High70,Lur70,Medsy70,Lpc70	0.045

P.M.P. of variables: South=0.6822, Central=0.0000, NEast=0.0073, High70=0.0000, Coll70=0.7941, Mfgs70=0.7547, Lur70=0.5969, Medsy70=0.4836, Lpc70=0.1855, Lpop70=0.4681

Table 24
MSA Population Growth and Education (Geweke's)

Variables	Model posterior probability
Central,NEast,Medsy70,Lpc70,Lpop70	0.049
Central,NEast,Medsy70,Lur70,Lpc70,Lpop70	0.041
Central,NEast,Medsy70,Lur70,Mfgs70,Lpc70,Lpop70	0.039
Central,NEast,Medsy70,Mfgs70,Lpc70,Lpop70	0.037
Central,NEast,Mfgs70,Lpc70	0.025
Central,NEast,Lur70,Mfgs70,Lpc70	0.024
Central,NEast,Medsy70,Lpop70	0.021

P.M.P. of variables: South=0.6921, Central=0.0226, NEast=0.1716, Medsy70=0.2849, High70=0.7335, Coll70=0.8224, Lur70=0.5314, Mfgs70=0.3629, Lpc70=0.2530, Lpop70=0.2902

Table 25
CITY Population Growth and Education (MBVS)

Variables	Model posterior probability
Central,NEast,Lur70	0.061
Central,NEast,Lpc70	0.050
Central,NEast,Lur70,Lpc70	0.040
South,Central,NEast,Lur70	0.029
South,Central,NEast,Lpc70	0.026

P.M.P. of variables: South=0.6319, Central=0.2145, NEast=0.3075, High70=0.8720, Coll70=0.8834, Mfgs70=0.9999, Lur70=0.4713, Medsy70=0.6523, Lpc70=0.4362, Lpop70=0.7411, Lpmed70=0.9623

Table 26
MSA population growth and education (MBVS)

Variables	Model posterior probability
Central,NEast	0.024
Central,NEast,Lpc70	0.024
Central,NEast,Lur70	0.021
Central,NEast,Lur70,Lpc70	0.020
Central,NEast,Medsy70	0.015
Central	0.015

P.M.P. of variables: South=0.5795, Central=0.3130, NEast=0.4797, Medsy70=0.5493, High70=0.9986, Coll70=0.9858, Lur70=0.5387, Mfgs70=0.9925, Lpc70=0.5039, Lpop70=0.5809

Table 27
CITY Population Growth, Education and Income

Variables included	R^2	*Adjusted* R^2	*AIC*	*BIC*	*SBC*
Lpop70,Mfgs70, Nonw70,Medic70 Central,NEast	0.602	0.568	-223.600	-219.700	-207.200
Lpop70,Nonw70, Medic70,Central NEast	0.590	0.561	-223.300	-220.100	-209.200
Lpop70,Pc70,Mfgs70, Nonw70,Medic70, Central,NEast	0.609	0.569	-223.000	-218.500	-204.200
Lpop70,Pc70,Nonw70 Medic70,Central, NEast	0.595	0.560	-222.300	-218.700	-205.900
Mfgs70,Nonw70, Medic70,Central, NEast	0.584	0.554	-222.200	-219.100	-208.100

Table 28
CITY Population Growth and Education with Race

Variables included	R^2	*Adjuted* R^2	*AIC*	*BIC*	*SBC*
Mfgs70,Nonw70,South Central,NEas	0.564	0.534	-218.600	- 215.100	-204.600
Mfgs70,Nonw70,Central, NEast	0.544	0.504	-217.200	-214.400	-205.500
Lpc70,Mfgs70,Nonw70, Central,NEast	0.554	0.528	-216.900	-213.700	-202.900
Medsy70,Mfgs70, Nonw70,Central, NEast	0.558	0.527	-217.500	-214.200	-203.400
Medsy70,Mfgs70,Nonw70, Central,NEast	0.571	0.517	-217.800	-213.800	-201.400

Table 29
CITY Population Growth and Education with Race (BMA)

Variables included	Model posterior probability
South,Central,Lur70,Medsy70,Lpc70 Lpop70,Lmedic70	0.361
South,Central,Mfgs70,Nonw70,Medsy70,Lpc70, Lpop70,Lmedic70	0.071
South,Central,Nonw70,Medsy70,Lpc70,Lpop70, Lmedic70	0.066
South,Nonw70,Lur70,Medsy70,Lpc70,Lpop70, Lmedic70	0.055
South,Central,NEast,Mfgs70,Nonw70,Medsy70, Lpc70,Lpop70,Lmedic70	0.049
Central,Nonw70,Lur70,Medsy70,Lpc70,Lpop70, Lmedic70	0.046

Posterior t-probability: South=0.0131, Central=0.0080, NEast=0.9640, Mfgs70=0.9744, Nonw70=0.4477, Lur70=0.4525, Medsy70=0.0026, Lpc70=0.9866, Lpop70=0.0000, Lmedic70=0.0087

Table 30
MSA Population Growth and Education with Race (BMA)

Variables	Model posterior probability
South,NEast,Mfgs70,Lpc70	0.057
South,Central,NEast,Lur70,Mfgs70	0.052
South,NEast,Lur70,Nonw70	0.048
South,Central,NEast,Lur70,Nonw70,	0.043
South,Central,NEast,Mfgs70	0.025
Central,NEast,Lur70,Nonw70,Lpop70	0.022

Posterior t-probability: South=0.4831, Cenral=0.8000, NEast=0.0165, Medsy70=0.7724, Lur70=0.4302, Mfgs70=0.2480, Lpc70=0.7630, Nonw70=0.2553, Lpop70=0.7889

Table 31
CITY Population Growth and Education with Race (Geweke's)

Variables	Model posterior probability
South,Central,NEast,Nonw70,Medsy70,Lpop70	0.086
South,Central,NEast,Nonw70,Medsy70,Lpc70,Lpop70	0.079
South,Central,NEast,Nonw70,Lur70,Medsy70,Lpop70	0.073
South,Central,NEast,Nonw70,Mesdsy70,Lpc70	0.060
South,Central,NEast,Nonw70,Lur70,Medsy70,Lpc70,Lpop70	0.052
South,Central,NEast,Nonw70,Medsy70	0.049

P.M.P. of variables: South=0.4267, Central=0.0000, NEast=0.0012, Mfgs70=0.0042, Nonw70=0.0005, Lur70=0.6681, Medsy70=0.1852, Lpc70=0.5033, Lpop70=0.7671, Lmedic70=0.5135

Table 32
MSA Population Growth and Education with Race (Geweke's)

Variables	Model posterior probability
Central,NEast,Medsy70,Lpc70,Lpop70	0.043
Central,NEast,Medsy70,Lur70,Lpc70,Lpop70	0.042
Central,NEast,Medsy70,Mfgs70,Lpc70,Lpop70	0.042
Central,NEast,Medsy70,Lur70,Mfgs70,Lpc70,Lpop70	0.040
Central,NEast,Mfgs70,Lpc70	0.025
Central,NEast,Medsy70,Lpop70	0.022

P.M.P. of variables: South=0.6789, Central=0.0214, NEast=0.1662, Medsy70=0.3115, Lur70=0.5329, Mfgs70=0.3833, Lpc70=0.2063, Nonw70=0.6853, Lpop70=0.3623

Table 33
CITY Population Growth and Education with Race (MBVS)

Variables	Model posterior probability
Central,NEast,Lur70	0.081
Central,NEast,Lpc70	0.069
Central,NEast,Lur70,Lpc70	0.055
South,Central,NEast,Lur70	0.039
South,Central,NEast,Lpc70	0.037
Central,NEast,Medsy70,Lpc70	0.033

P.M.P. of variables: South=0.6344, Central=0.1799, NEast=0.2458, Mfgs70=0.9997, Nonw70=0.9933, Lur70=0.4488, Medsy70=0.6069, Lpc70=0.4204, Lpop70=0.7242, Lpmdic70=0.9897

Table 34
MSA Population Growth and Rducation with Race (MBVS)

Variables	Model posterior probability
Central,NEast	0.023
Central,NEast	0.022
Central,NEast,Lur70	0.020
Central,NEast,Lur70,Lpc70	0.019
Central,NEast,Lpop70	0.016
Central,NEast,Lpc70,Lpop70	0.016

P.M.P. of variables: South=0.5786, Central=0.2985, NEast=0.4761, Mfgs70=0.9918, Nonw70=0.9993, Lur70=0.5387, Medsy70=0.5938, Lpc70=0.5035, Lpop70=0.5236

Table 35
CITY Population Growth and Inequality

Variables included	R^2	*Adjusted* R^2	*AIC*	*BIC*	*SBC*
Lpop70,Lpc70,Mfgs70, Incle70,Lmedic70, Central,NEast	0.613	0.573	-223.800	-219.300	-205.100
Lpop70,Mfgs70,Incle70, South,Central	0.588	0.567	-222.900	-219.900	-208.900
Lpop70,Mfgs70,Incle70, Edle70,Lmedic70,Central, NEast	0.610	0.571	-223.100	-218.800	-204.400
Lpop70,Mfgs70,Incle70, Edle70,Central,NEast	0.601	0.561	-223.400	-219.700	–207.000
Lpop70,Mfgs70,Lpc70, Incle70,Edle70,Lmedic70, Central,NEast	0.62	0.578	-223.700	-218.400	-202.600

Table 36
MSA Population Growth and Inequality

Variables included	R^2	*Adjusted* R^2	*AIC*	*BIC*	*SBC*
Lpop70,Central,NEast	0.283	0.253	-182.800	-180.200	-173.500
Lpop70,Mfgs70,Central, NEast	0.300	0.260	-182.500	-179.600	-171.000
Lpop70,Mfgs70,Central	0.274	0.243	-181.800	-179.400	-172.600
Lpop70,Lur70,Central, NEast	0.293	0.252	-181.800	-179.000	-170.200
Lpop70,Lur70,Mfgs70, Medsy70,Central,NEast	0.311	0.261	-181.300	-178.500	-167.900

Table 37
CITY Population Growth and Inequality (BMA)

Variables included	Model posterior probability
South,Mfgs70,Incle70, Lpop70,Edle70,Lmedic70	0.100
South,Mfgs70,Incle70,Incla70,Lpop70, Edle70,Medic70	0.050
South,NEast,Mfgs70,Medsy70,Incle70, Lpop70,Edle70,Medic70	0.047
South,Mfgs70,Lur70,Medsy70,Incle70,Lpc70, Lpop70,Edle70,Lmedic70	0.026
South,NEast,Mfgs70,Medsy70,Lpop70, Edle70,Lmedic70	0.019
South,Mfgs70,Medsy70,Incle70,Incla70,Lpop70, Edle70,Lmedic70	0.019

Posterior t-probability: South=0.0542, Central=0.7979, NEast=0.6767, Mfgs70=0.0150, Lur70=0.9581, Medsy70=0.3896, Incle70=0.0817, Inlca70=0.6583, Lpc70=0.3913,Lpop70=0.8104, Edle70=0.0000, Lmedic70=0.0006

Table 38
MSA Population Growth and Inequality (BMA)

Variables	Model posterior probability
South,NEast,Incle70	0.037
NEast,Incle70	0.036
NEast,Lur70,Lpc70	0.028
NEast,Lur70,Incle70	0.026
South,NEast,Lur70,Lpc70	0.017
South,NEast,Lur70,Incle70	0.015

Posterior t-probability: South=0.7372, Central=0.8629, NEast=0.0360, Medsy70=0.9111, Lur70=0.7540, Mfgs70=0.7061, Incle70=0.2627, Incla70=0.8237, Lpc70=0.2480, Lpop70=0.7241

Table 39
CITY Population Growth and Inequality (Geweke's)

Variables	Model posterior probability
Central,NEast,Medsy70,Incle70,Lpop70	0.581
Central,NEast,Medsy70,Incle70	0.082
Central,NEast,Medsy70,Incle70,Edle70,Lpop70	0.068
Central,NEast,Lur70,Medsy70,Incle70,Lpop70	0.066
Central,NEast,Medsy70,Incle70,Lpc70,Lpop70	0.038
South,Central,NEast,Medsy70,Incle70,Lpop70	0.019
Central,NEast,Medsy70,Incle70,Incla70,Lpop70	0.016
Central,NEast,Medsy70,Incle70,Lmedic70,Lpop70	0.015
Central,NEast,Medsy70,Incle70,Lpop70	0.014
Central,NEast,Coll70,Medsy70,Incle70,Lpop70	0.013

P.M.P. of variables: South=0.7580, Central=0.0000, NEast=0.0000, Coll70=0.8190, Mfgs70=0.1475, Lur70=0.4652, Medsy70=0.0000, Incle70=0.0000, Incla70=0.8975, Lpc70=0.6344, Edle70=0.3098, Lmedic70=0.5273, Lpop70=0.0790

Table 40
MSA [opulation Growth and Inequality (Geweke's)

Variables	Model posterior probability
Central,NEast,Medsy70,Lur70,Lpc70,Lpop70	0.044
Cenral,NEast,Medsy70,Lur70,Mfgs70,Lpc70,Lpop70	0.044
Central,NEast,Medsy70,Lpc70,Lpop70	0.039
Central,NEast,Medsy70,Mfgs70,Lpc70,Lpop70	0.034
Cenral,NEast,Mfgs70,Lpc70	0.022
Central,NEast,Mfgs70,Lpc70,Lpop70	0.021
Central,NEast,Lpc70,Lpop70	0.020

P.M.P. of variables: South=0.7165, Central=0.0254, NEast=0.2209, Medsy70=0.1850, Lur70=0.5281, Mfgs70=0.3688, Incle70=0.7951, Incla70=0.7320, Lpc70=0.3423, Lpop70=0.2569

Table 41
CITY Population Growth and Inequality (MBVS)

Variables	Model posterior probability
Central,NEast,Medsy70	0.025
Central,NEast,Medsy70,Lmedic70	0.025
Central,NEast,Medsy70,Lpc70	0.025
Central,NEast,Medsy70,Lpc70,Lmedic70	0.024
Central,Medsy70	0.023
Central,Medsy70,Lmedic70	0.023
Central,Medsy70,Lpc70	0.022

P.M.P. of variables: South=0.6288, Central=0.2356, NEast=0.4813, Coll70=0.9779, Mfgs70=0.9999, Lur70=0.5277, Medsy70=0.0879, Incle70=0.9693, Incla70=0.9780, Lpc70=0.5012, Lpop70=0.8504, Lmedic70=0.5000

Table 42
MSA Population Growth and Inequality (MBVS)

Variables	Model posterior probability
Central,NEast	0.023
Central,NEast,Lpc70	0.022
Central,,NEast,Lur70	0.020
Central,NEast,Lur70,Lpc70	0.019
Central,NEast,Medsy70	0.015
Central	0.022

P.M.P. of variables: South=0.5811, Central=0.3145, NEast=0.4812, Medsy70=0.5936, Lur70=0.5390, Mfgs70=0.9932, Inlcle70=0.9751, Incla70=0.9585, Lpc70=0.5037, Lpop70=0.5859

Table 43
CITY Population Growth and Social Characteristics

Variables included	R^2	*Adjusted* R^2	*AIC*	*BIC*	*SBC*
Lpop70,Mfgs70,Age70, South,Central,NEast, Seg70,Weseg70	0.875	0.857	-261.300	-255.200	-242.000
Lpop70,Mfgs70,Age70, Nonw70,South,Central, NEast	0.871	0.854	-261.100	-256.100	-244.000
Lpop70,Mfgs70,Age70, Weseg70,South,Central, NEast	0.870	0.853	-260.500	-255.700	-243.400
Lpop70,Mfgs70,Age70, Nonw70,Lur70,South, Central,NEast	0.874	0.856	-260.800	-254.800	-241.500
Lpop70,Mfgs70,Age70, Nonw70,Weseg70,South, Central,NEast	0.872	0.853	-259.800	-254.200	-240.600

Table 44
MSA Population Growth and Social Characteristics

Variables included	R^2	*Adjusted* R^2	*AIC*	*BIC*	*SBC*
Lpop70,Lpc70,Nonw70, Mfgs70,Medsy70, Central, NEast	0.480	0.413	-159.600	-154.400	-142.400
Lpop70,Lpc70,Nonw70, Lur70,Mfgs70,Medsy70	0.444	0.385	-157.500	-153.700	-142.500
Lpop70,Lpc70,Nonw70, Mfgs70,Medsy70, Central	0.451	0.392	-158.200	-154.300	-143.200
Lpop70,Lpc70, Nonw70, Mfgs70,Medsy70	0.433	0.383	-158.100	-154.900	-145.300
Lpop70,Lpc70,Nonw70, Mfgs70,Medsy70,South Central,NEast	0.484	0.408	-158.100	-152.400	-138.800

Table 45
CITY Population and Social Characteristics (BMA)

Variables included	Model posterior probability
South,Central,NEast,Lpc70,Lpop70,Seg70,Weseg70	0.040
South,Central,NEast,Lpc70,Lpop70,Seg70,Weseg70	0.039
South,Central,Lpc70,Lpop70,Seg70,Weseg70	0.023
South,Central,NEast,Age70,Nonw70,Lpc70,Lpop70, Seg70,Weseg70	0.020
South,Central,Mfgs70,Lpc70,Lpop70,Seg70,Weseg70	0.019
South,Central,Age70,Lur70,Lpc70,Lpop70,Seg70, Weseg70	0.019

For Table 24, only has 63 cities because the segregation index is not available for all 77 cities.
Posterior t-probability: South=0.0000, Central=0.0069, NEast=0.7457, Age70=0.9838,Mfgs70=0.6653, Nonw70=0.8157, Lur70=0.6446, Medsy70=0.8908, Lpc70=0.0001, Seg70=0.0456, Lpop70=0.0000, Weseg70=0.0065

Table 46
MSA Population Growth and Social Characteristics (BMA)

Variables	Model posterior probability
South,NEast,Mfgs70,Age70,Weseg70	0.008
South,NEast,Lur70,Mfgs70,Age70,Weseg70	0.008
South,NEast,Mfgs70,Age70,Lpop70	0.007
South,NEast,Lur70,Mfgs70,Age70,Lpop70	0.006
South,Central,NEast,Lur70,Mfgs70,Age70,Lpop70	0.001

Posterior t-probability: South=0.1318, Central=0.9235, NEast=0.0036, Medsy70=0.6819, Lur70=0.7251, Mfgs70=0.1587, Lpc70=0.8816, Age70=0.0271, Nonw70=0.9878, Seg70=0.9221, Lpop70=0.4811, Weseg70=0.8338

Table 47
CITY Population Growth and Social Characteristics (Geweke's)

Variables	Model posterior probability
South,Central,NEast,Age70,Mfgs70,Medsy70,Lpc70, Lpop70,Seg70,Weseg70	0.105
South,Central,NEast,Age70,Medsy70,Lpc70,Seg70,Weseg70	0.067
South,Central,NEast,Age70,Mfgs70,Medsy70,Lpc70,Weseg70	0.060
South,Central,NEast,Age70,Mfgs70,Lur70,Medsy70,Lpc70, Lpop70,Seg70,Weseg70	0.054
South,Central,NEast,Age70,Medsy70,Lpc70,Lpop70,Weseg70	0.046
South,Central,NEast,Age70,Medsy70,Lpc70,Lpop70,Seg70,Weseg70	0.041
South,Central, NEast,Age70,Mfgs70,Lur70,Medsy70,Lpc70,Seg70, Weseg70	0.039
South,Central,NEast,Age70,Mfgs70,Medsy70,Lpc70,Lpop70,Weseg70	0.030
South,Central,NEast,Age70,Mfgs70,Lur70,Medsy70,Lpc70,Weseg70	0.027

For this table, only 63 cities have dissimilarity index which is used to measure the segregation index between white and nonwhite population.

P.M.P. of variables: South=0.0868, Central=0.0000, NEast=0.0003, Age70=0.0000, Mfgs70=0.3320 Nonw70=0.8093, Lur70=0.6321, Medsy70=0.0000, Lpc70=0.2106, Lpop70=0.4124, Seg70=0.4222, Weseg70=0.0873

Table 48
MSA Population Growth and Social Characteristics (Geweke's)

Variables	Model posterior probability
Central,NEast,Medsy70,Mfgs70,Lpc70,Lpop70	0.026
Central,NEast,Medsy70,Lpc70,Lpop70	0.015
Central,NEast,Medsy70,Mfgs70,Lpc70,Lpop70,Seg70	0.012
Central,NEast,Mfgs70,Lpc70	0.011
South,Central,NEast,Medsy70,Mfgs70,Lpc70,Lpop70	0.011
Central,NEast,Medsy70,Mfgs70,Lpc70	0.010
Central,NEast,Medsy70,Mfgs70,Lpop70	0.009
Central,NEast,Medsy70,Mfgs70,Lpc70,Nonw70,Lpop70	0.008
Central,NEast,Medsy70,Mfgs70	0.008

P.M.P. of variables: South=0.4854, Central=0.1878, NEast=0.2937, Medsy70=0.2479, Lur70=0.5958, Mfgs70=0.1031, Lpc70=0.3334, Age70=0.8090, Nonw70=0.6780, Lpop70=0.5475, Seg70=0.6223, Weseg70=0.8218

Table 49
CITY Population and Social Characteristics (MBVS)

Variables	Model posterior probability
Central,NEast,Medsy70,Lpc70	0.012
Central,NEast,Medsy70,Lpc70,Seg70	0.012
Central,NEast,Medsy70	0.012
Central,NEast,Medsy70,Seg70	0.012
Central,Medsy70,Lpc70	0.010
Central,Medsy70,Lpc70,Seg70	0.010
Central,Medsy70	0.010

For this table, only use 63 cities to analyze which have available segregation index.
P.M.P. of variables: South=0.6087, Central=0.3284, NEast=0.4959, Age70=0.6556, Mfgs70=0.9992, Nonw70=0.9996, Lur70=0.5337, Medsy70=0.2875, Lpc70=0.4999, Seg70=0.5003, Lpop70=0.6376, Weseg70=0.9971

Table 50
MSA Population Growth and Social Characteristics (MBVS)

Variables	Model posterior probability
Central,Lpop70	0.009
Central,Lpc70,Lpop70	0.009
Central,Lpop70,Seg70	0.008
Central,Lpc70,Lpop70,Seg70	0.008
Central,Lur70,Lpop70	0.008
Central,Lur70,Lpc70,Lpop70	0.008
Central,NEast,Lpop70	0.007
Central,NEast,Lpc70,Lpop70	0.007
Central,Lur70,Lpop70,Seg70	0.007

P.M.P. of variables: South=0.6057, Central=0.4320, NEast=0.5184, Medsy70=0.5920, Lur70=0.5325, Mfgs70=0.9934, Lpc70=0.5007, Age70=0.9736, Nonw70=0.9997, Lpop70=0.4115, Seg70=0.5090, Weseg70=0.9978

Table 51
CITY Population Growth and Government

Variables included	R^2	*Adjusted* R^2	*AIC*	*BIC*	*SBC*
Lpop70,Medsy70,Mfgs70, Expo70,Ldebt70,Central, NEast	0.579	0.494	-217.300	-212.000	-198.500
Lpop70,Medsy70,Mfgs70, Expo70,South,Central, NEast	0.586	0.486	-216.500	-210.500	-195.400
Lpop70,Medsy70,Expo70, Ldebt70,Central,NEast	0.563	0.484	-216.400	-212.200	-200.000
Lpop70,Medsy70,Mfgs70, Ldebt70,Central,NEast	0.562	0.476	-216.200	-212.000	-199.800
Lpop70,Medsy70,Expo70, Ldebt70,South,Central, NEast	0.569	0.473	-215.500	-210.700	-196.800

Table 52
MSA Population Growth and Government

Variables included	R^2	*Adjusted* R^2	*AIC*	*BIC*	*SBC*
CentrL,NEast,Lur70, Lgvpc70,Pctax70, Exedu70,Lpc70,Lpop70	0.436	0.368	-190.800	-185.200	-170.000
Lpop70,Lur70, Lgvpc70,Pctax70, Exedu70,Central,NEast	0.421	0.361	-190.800	-186.100	-172.300
Lpop70,Lur70,Lvpc70, Pctax70,Exedu70,South, Central, NEast	0.429	0.361	-190.000	-184.600	-169.100
Lpop70,Lur70,Lgvpc70, Lpc70, Exedu70,Central, NEast	0.414	0.353	-189.900	-185.400	-171.400
Lpop70,Lur70,Lgvpc70, Exedu70,Pctax70,Exhea70, Central,NEast	0.426	0.357	-189.500	-184.200	-168.600

Table 53
CITY Population Growth and Government (BMA)

Variables	Model posterior probability
South,NEast,Mfgs70,Medsy70,Igr70, Exhwy70,Lpc70,Lpop70	0.015
South,NEast,Mfgs70,Lur70,Medsy70,Igr70,Lpcex70, Exhwy70,Lpc70,Lpop70	0.012
South,NEast,Mfgs70,Lur70,Igr70,Exhwy70, Lpc70,Lpop70	0.011
South,NEast,Mfgs70,Exhwy70,Lpc70,Lpop70	0.010

Posterior t-probability: South=0.2473, Central=0.8578, NEast=0.0066, Mfgs70=0.7053, Lur70=0.8272, Medsy70=0.8353, Lgvpc70=0.8567, Igr70=0.6688, Lpcex70=0.7982, Exhwy70=0.1808, Expo70=0.9565, Exss70=0.8400, Ldebt70=0.9724, Lpc70=0.0002, Lpop70=0.1239

Table 54
MSA Population Growth and Government (BMA)

Variables	Model posterior probability
Central,NEast,Medsy70,Mfgs70,Igr70,Pctax70, Lpcex70,Exhwy70	0.015
Central,NEast,Mfgs70,Igr70,Pctax70,Lpcex70, Exhwy70	0.014
Central,NEast,Medsy70,Mfgs70,Igr70,Pctax70, Exhwy70	0.013
Central,NEast,Mfgs70,Igr70,Exhwy70	0.013
South,NEast,Medsy70,Mfgs70,Igr70,Pctax70, Lpcex70,Exhwy70	0.012
Central,NEast,Mfgs70,Igr70,Pctax70,Exhwy70	0.012
NEast,Mfgs70,Igr70,Exhwy70	0.012
NEast,Medsy70,Mfgs70,Igr70,Pctax70,Lpcex70, Exhwy70	0.011

Posterior t-probability: South=0.7348, Central=0.5604, NEast=0.0844, Medsy70=0.6101, Lur70=0.9600, Mfgs70=0.1350, Lgvpc70=0.8439, Igr70=0.1530, Pctax70=0.6853, Lpcex70=0.7709, Exedu70=0.9048, Exhwy70=0.0152, Exheal70=0.9657, Lpc70=0.6802, Lpop70=0.9598

Table 55
CITY Population and Government (Geweke's)

Variables	Model posterior probability
South,Central,NEast,Mfgs70,Medsy70,Lpcex70,Ldebt70	0.012
South,Central,NEast,Mfgs70,Medsy70,Ldebt70,Lpc70	0.011
South,Central,NEast,Mfgs70,Medsy70,Ldebt70	0.010
South,Central,NEast,Mfgs70,Medsy70,Lpcex70,Expo70,Ldebt70,Lpc70, Lpop70	0.010
South,Central,NEast,Mfgs70,Medsy70,Lgvpc70,Lpcex70,Expo70,Ldebt70, Lpop70	0.009
South,Central,NEast,Mfgs70,Medsy70,Ldebt70,Lpop70	0.009
South,Central,NEast,Mfgs70,Medsy70,Lgvpc70,Ldebt70	0.009
South,Central,NEast,Mfgs70,Mesdsy70,Ldebt70,Lpc70	0.009
South,Central,NEast,Mfgs70,Medsy70,Lpcex70,Ldebt70,Lpc70	0.009
South,Central,NEast,Mfgs70,Mesdsy70,Lgvpc70,Ldebt70,Lpc70	0.008
South,Central,NEast,Mfgs70,Medsy70,Lpcex70,Expo70,Ldebt70,Lpop70	0.008
South,Central,NEast,Mfgs70,Medsy70,Lgvpc70,Lpcex70,Expo70,Ldebt70 Lpc70,Lpop70	0.008

P.M.P. of variables: South=0.3295, Central=0.0000, NEast=0.0133, Mfgs70=0.0672, Lur70=0.6082, Medsy70=0.0000, Lgvpc70=0.5490, Igr70=0.9125, Lpcex70=0.4925, Exhwy70=0.7945, Expo70=0.5408, Exss70=0.9010, Ldebt70=0.1014, Lpc70=0.5284, Lpop70=0.5050

Table 56
MSA Population and Government (Geweke's)

Variables	Model posterior probability
Central,NEast,Mfgs70,Lpcex70,Lpc70,Lpop70	0.022
Central,NEast,Medsy70,Mfgs70,Lpcex70,Lpc70,Lpop70	0.019
Central,NEast,Medsy70,Lpcex70,Lpc70,Lpop70	0.018
Central,NEast,Medsy70,Lpcex70,Lpc70,Lpop70	0.015
Central,NEast,Medsy70,Mfgs70,Lpcex70,Lpop70	0.012
Central,NEast,Medsy70,Lur70,Mfgs70,Lpcex70,Lpop70	0.012
Central,NEast,Medsy70,Lur70,Lpcex70,Lpc70,Lpop70	0.011
Central,NEast,Medsy70,Lpcex70,Lpop70	0.010
Central,NEast,Medsy70,Lur70,Mfgs70,Lpcex70,Lpc70,Lpop70	0.010
Central,NEast,Lgvpc70,Lpcex70,Lpc70,Lpop70	0.009
Central,NEast,Lpcex70,Lpc70,Lpop70	0.009
Central,Mfgs70,Lpcex70,Lpc70,Lpop70	0.008

P.M.P. of variables: South=0.7342, Central=0.0134, NEast=0.1938, Medsy70=0.2777, Lur70=0.5645, Mfgs70=0.3455, Lgvpc70=0.6477, Igr70=0.9307, Pctax70=0.8625, Lpcex70=0.1859, Exedu70=0.7600, Exhwy70=0.8971, Exheal70=0.8570, Lpc70=0.2761, Lpop70=0.1352

Table 57
CITY Population Growth and Government (MBVS)

Variables	Model posterior probability
Central,NEast,Mfgs70	0.022
Central,NEast,Mfgs70,Lpc70	0.022
Central,Mfgs70	0.021
Central,Mfgs70,Lpc70	0.021
Central,NEast,Lur70,Mfgs70	0.019
Central,NEast,Lur70,Mfgs70,Lpc70	0.019
Central,Lur70,Mfgs70	0.018
Central,Lur70,Mfgs70,Lpc70	0.018
South,Central,NEast,Mfgs70	0.012
South,Central,NEast,Mfgs70	0.011

P.M.P. of variables: South=0.6323, Central=0.2614, NEast=0.5050, Mfgs70=1.0000, Lur70=0.5287, Medsy70=0.1450, Lgvpc70=0.7807, Igr70=0.9999, Lpcex70=0.6893,Exhwy70=0.9974, Expo70=0.9993, Exss70=0.9987, Ldebt70=0.8402, Lpc70=0.5012, Lpop70=0.8454

Table 58
MSA Population Growth and Government (MBVS)

Variables	Model posterior probability
Central,NEast	0.004
Central	0.003
Central,NEast,Medsy70	0.002
Central,NEast,Lpc70	0.002
South,Central,NEast,	0.002
South,Central	0.002
Central,Medsy70	0.002
Central,NEast,Lgvpc70	0.002
Central,Lpc70	0.002
Central,Lgvpc70	0.002

P.M.P. of variables: South=0.5886, Central=0.3739, NEast=0.5062, Medsy70=0.5957, Lur70=0.9488, Mfgs70=0.9832, Lgvpc70=0.6255, Igr70=0.9999, Pctax70=0.9980, Lpcex70=0.9141, Exedu70=0.9864, Exhwy70=0.9992, Exheal70=0.9993, Lpc70=0.6131, Lpop70=0.9135

Table 59
CITY Population Growth and all Variables

Variables included	R^2	*Adjutsed* R^2	*AIC*	*BIC*	*SBC*
South,Central,NEast, Age70,Nonw70,Lgvpc70, Igr70,Lpcex70,Exhwy70, Ldebt70,Lpc70,Lmedic70	0.865	0.804	-294.700	-283.200	-264.200
South,Central,NEast, Age70,Nonw70,Lgvpc70, Igr70,Lpcex70,Exhwy70 Lpc70,Lmedic70	0.860	0.805	-293.800	-284.200	-265.700
South,Central,NEast, Age70,Mfgs70,Nonw70, Lgvpc70,Igr70,Lpcex70, Exhwy70,Lpc70,Lmedic70	0.864	0.805	-294.200	-282.900	-263.700
South,Central,NEast, Age70,Mfgs70,Nonw70, Incla70,Lgvpc70,Lpcex70 Ldebt70,Edle70	0.859	0.805	-293.500	-284.000	-265.400
South,Central,NEast, Age70,Nonw70,Lgvpc70, Igr70,Lpcex70,Exhwy70, Lmedic70	0.859	0.804	-293.400	-284.000	-265.300

Table 60
MSA Population Growth and all Variables

Variables included	R^2	*Adjusted* R^2	*AIC*	*BIC*	*SBC*
Central,NEast,Lur70, Lgvpc70,Pctax70,Lpop70	0.424	0.364	-191.200	-185.500	-172.700
Central,NEast,Lur70, Mfgs70,Lgvpc70, Lpop70	0.408	0.356	-191.200	-184.600	-175.000
Central,NEast,Lur70, Mfgs70,Lgvpc70,Pctax70, Exedu70,Lpop70	0.438	0.370	-191.000	-184.400	-170.100
Central,NEast,Lur70, Lgvcp70,Pctax70,Exedu70, Lpop70	0.436	0.368	-190.800	-184.200	-170.000
Central,NEast,Lur70, Mfgs70,Lgvpc70, Lpc70,Lpop70	0.421	0.361	-190.800	-185.300	-172.300
Central,NEast,Lur70, Lgvpc70,Pctax70, Exedu70,Lpop70	0.421	0.361	-190.800	-185.300	-172.300

Table 61
CITY Population Growth and all Variables (BMA)

Variables	Model posterior probability
South,Central,NEast,Coll70,Age70,Nonw70,Lur70, Medsy70,Incla70,Igr70,Expo70,Ldebt70,Lpop70	0.020
South,Central,NEast,Coll70,Age70,Nonw70, Lur70,Medsy70,Incla70,Lgvpc70,Igr70,Expo70, Ldebt70,Lpop70	0.017
South,Central,NEast,Coll70,Age70,Nonw70,Lur70, Medsy70,Incla70,Lgvpc70,Expo70,Ldebt70,Lpop70	0.011
South,Central,NEast,Coll70,Age70,Mfgs70,Nonw70, Nonw70,Lur70,Medsy70,Incla70,Igr70,Expo70, Ldebt70,Lpop70	0.011
South,Central,NEast,Coll70,Age70,Mfgs70,Nonw70, Lur70,Medsy70,Incla70,Lgvpc70,Igr70,Ldebt70,Lpop70	0.010
South,Central,NEast,Coll70,Age70,Nonw70,Lur70,Medsy70, Incla70,Lgvpc70,Igr70,Ldebt70,Lpop70	0.010
South,Central,NEast,Coll70,Age70,Mfgs70,Nonw70,Lur70, Medsy70,Incla70,Igr70,Ldebt70,Lpop70	0.010

Table 62
CITY Population Growth and all Variables (Geweke's)

Variables	Model posterior probability
South,Central,NEast,High70,Age70,Nonw70,Lgvpc70,Lpcex70,Ldebt70, Lpc70,Edle70	0.020
South,Central,NEast,High70,Age70,Nonw70,Lgvpc70,Ldebt70, Lpc70,Edle70,Lpop70	0.018
South,Central,NEast,High70,Age70,Nonw70,Medsy70,Lgvpc70,Lpcex70, Ldebt70,Lpc70,Edle70	0.014
South,Central,NEast,High70,Age70,Nonw70,Lur70,Lgvpc70,Lpcex70, Ldebt70,Lpc70,Edle70	0.010
South,Central,NEast,High70,Age70,Nonw70,Medsy70,Incla70,Lgvpc70, Lpcex70,Ldebt70,Lpc70,Edle70,Lpop70	0.010
South,Central,NEast,High70,Age70,Nonw70,Lgvpc70,Lpcex70,Ldebt70, Lpc70,Edle70,	0.009
South,Central,NEast,High70,Age70,Nonw70,Igr70,Lpcex70,Ldebt70 Lpc70,Edle70	0.009
South,Central,NEast,High70,Age70,Lur70,Medsy70,Lgvpc70,Lpcex70, Ldebt70,Lpc70,Edle70,Lpop70	0.009
South,Central,NEast,High70,Age70,Nonw70,Lgvpc70,Igr70,Lpcex70, Ldebt70,Lpc70,Edle70,Lpop70	0.008
South,Central,NEast,High70,Age70,Nonw70,Lpcex70, Ldebt70, Lpc70,Edle70,Lpop70	0.007
South,Central,NEast,High70,Age70,Nonw70,Medsy70,Lgvpc70,Igr70, Lpcex70,Ldebt70,Lpc70,Edle70	0.007
South,Central,NEast,High70,Age70,Nonw70,Lgvpc70,Lpcex70, Ldebt70,Edle70	0.006
South,Central,NEast,High70,Age70,Nonw70,Lur70,Medsy70,Lgvpc70, Lpcex70,Ldebt70,Lpc70,Edle70	0.006
South,Central,NEast,High70,Age70,Nonw70,Lgvpc70,Lpcex70,Ldebt70, Edle70,Lpop70	0.006
South,Cenral,NEast,High70,Age70,Nonw70,Lgvpc70,Lpcex70,Ldebt70, Lpc70,Edle70	0.005
South,Central,NEast,High70,Age70,Nonw70,Medsy70,Lgvpc70,Lpcex70, Ldebt70,Lpc70,Edle70,Lpop70	0.005
South,Central,NEast,Hgih70,Age70,Nonw70,Lgvpc70,Lpcex70,Ldebt70 Lpc70,Edle70,Lpop70	0.005
South,Central,NEast,High70,Age70,Nonw70,Lur70,Lgvpc70,Igr70, Lpcex70,Ldebt70,Lpc70,Edle70,Lpop70	0.005
South,Central,NEast,High70,Age70,Nonw70,Lur70,Medsy70,Lgvpc70, Lpcex70,Ldebt70,Lpc70,Edle70,Lpop70	0.004
South,Central,High70,Age70,Nonw70,Medsy70,Lgvpc70,Lpcex70, Ldebt70,Edle70,Lpop70	0.004

Table 63
MSA Population Growth and all Variables (Geweke's)

Variables	Model posterior probability
Central,NEast,Medsy70,Lur70,Incle70,Lpcex70,Exedu70,Nonw70,Lpop70	0.002
Central,NEast,Medsy70,Lpcex70,Lpop70	0.002
Central,NEast,Medsy70,Lur70,Mfgs70,Incle70,Lpcex70,Lpop70	0.001
Central,NEast,Medsy70,Lur70,Lpcex70,Lpop70	0.001
Central,NEast,Medsy70,Lur70,Mfgs70,Lpcex70,Lpop70	0.001
Central,NEast,Medsy70,Mfgs70,Lpcex70,Lpop70	0.001
Central,NEast,Medsy70,Lur70,Mfgs70,Lgvpc70,Lpcex70,Lpop70	0.001
Central,NEast,Medsy70,Incle70,Lpcex70,Exedu70,Lpc70,Nonw70,Lpop70	0.001
Central,NEast,Medsy70,High70,Lur70,Mfgs70,Lpcex70,Nonw70	0.001
Central,NEast,Medsy70,Lur70,Incle70,Lgvpc70,Lpcex70,Exedu70, Lpc70,Age70,Nonw70,Lpop70	0.001
Central,NEast,Medsy70,Mfgs70,Lpcex70,Lpc70,Lpop70	0.001
Central,NEast,Mfgs70,Lpcex70,Lpc70,Lpop70	0.001
Central,NEast,High70,Lur70,Lpcex70,Exedu70,Lpc70,Age70,Nonw70	0.001
Central,NEast,Medsy70,Lur70,Lpop70	0.001
Central,NEast,Medsy70,Mfgs70,Lpop70	0.001
Central,NEast,Medsy70,Lgvpc70,Lpcex70,Lpc70,Lpop70	0.001
Central,NEast,Medsy70,Lgvpc70,Lpcex70,Lpc70,Lpop70	0.001
Central,NEast,Medsy70,Lgvpc70,Lpcex70,Lpop70	0.001

For this table, including all variables except the segregation index.

Table 64
CITY Population Growth and all Variables (MBVS)

Variables	Model posterior probability
Central,Medsy70	0.003
Central,Medsy70,Lmedic70	0.003
Central,Medsy70,Lpc70	0.003
Central,Medsy70,Lpc70,Lmedic70	0.003
Age70	0.003
Age70,Lpc70	0.003
Age70,Lmedic70	0.003
Age70,Lpc70,Lmedic70	0.003
Central,NEast,Medsy70	0.003
Central,NEast,Medsy70,Lmedic70	0.003
Central,NEast,Medsy70,Lpc70	0.003
Central,NEast,Medsy70,Lpc70	0.003
Central,Lur70,Meddsy70	0.003
Central,Lur70,Medsy70,Lmedic70	0.003
Central,Lur70,Medsy70,Lpc70	0.003
Central,Lur70,Medsy70,Lpc70,Lmedic70	0.003
Age70,Lur70	0.003
Age70,Lur70,Lmedic70	0.003
Age70,Lur70,Lpc70,	0.003
Age70,Lur70,Lpc70,Lmedic70	0.003

Table 65
MSA Population Growth and all Variables (MBVS)

Variables	Model posterior probability
Central,NEast	0.014
Central,NEast,Lpc70	0.014
Central	0.013
Central,Lpc70	0.013
Central,NEast,Lur70	0.012
Central,NEast,Lur70,Lpc70	0.012
Central,Lur70	0.011
Central,Lur70,Lpc70	0.011
Null model	0.010
Lpc70	0.009
Central,NEast,Medsy70	0.009
Central,NEast,Medsy70,Lpc70	0.009
Central,Medsy70	0.008
Central,Medsy70,Lpc70	0.008
NEast	0.008

Table 66
Posterior Marginal Probability of Three Bayesian Methods
(For CITY, all Variables)

Variable	BMA [1]	Geweke [2]	MBVS
South	0.096	0.135	0.638
Central	0.988	0.000	0.446
NEast	0.000	0.089	0.556
High70	0.832	0.156	0.996
Coll70	0.000	0.786	0.991
Age70	0.271	0.000	0.410
Mfgs70	0.882	0.948	1.000
Nonw70	0.210	0.000	1.000
Lur70	0.233	0.610	0.530
Medsy70	0.000	0.563	0.504
Incle70	0.953	0.801	0.986
Incla70	0.000	0.746	0.988
Lgvpc70	0.743	0.245	0.786
Igr70	0.514	0.922	1.000
Lpcex70	0.941	0.029	0.716
Exhwy70	0.932	0.914	0.999
Expo70	0.373	0.895	1.000
Exss70	0.948	0.890	0.999
Ldebt70	0.086	0.083	0.855
Lpc70	0.991	0.328	0.501
Lpop70	0.842	0.585	0.158
Edle70	0.878	0.159	0.986
Lmedic70	0.885	0.430	0.500

1: For BMA, is the posterior t-probability. As the usual t-probability, the small it is, the higher probability this variable is in the model.
2: For Geweke, is the posterior marginal probability that coefficient equal to zero or the posterior probability that this variable is not in the model.

Table 67
Posterior Marginal Probability of Three Bayesian MethodS
(For MSA, all Variables)

Variable	BMA	Geweke	MBVS
South	0.672	0.710	0.602
Central	0.714	0.021	0.431
NEast	0.967	0.289	0.532
Medsy70	0.393	0.304	0.596
High70	0.127	0.707	0.999
Coll70	0.966	0.779	0.989
Lur70	0.734	0.522	0.543
Mfgs70	0.554	0.468	0.997
Incle70	0.997	0.574	0.980
Incla70	0.191	0.691	0.965
Lgvpc70	0.870	0.628	0.852
Igr70	0.034	0.927	1.000
Pctax70	0.462	0.836	1.000
Lpcex70	0.863	0.209	0.807
Exedu70	0.935	0.678	0.999
Exhwy70	0.325	0.892	0.994
Exheal70	0.341	0.863	0.997
Lpc70	0.012	0.356	0.501
Age70	0.982	0.703	0.979
Nonw70	0.943	0.574	1.000
Lpop70	0.997	0.323	0.717

Table 68
Correlation Coefficients for Three Bayesian Methods (CITY, all Variables)
(Correlation Coefficients of P.M.P. of Three Bayesian Methods)

Pearson Correlation Coefficients / Prob > |R| under Ho: Rho=0 / N = 23

	B	G	M
B	1.00000 0.0	0.09794 0.6566	0.00908 0.9672
G	0.09794 0.6566	1.00000 0.0	0.52806 0.0096
M	0.00908 0.9672	0.52806 0.0096	1.00000 0.0

Table 69
Correlation Coefficients for three Bayesian methods (MSA, all variables)
(Correlation coefficients of P.M.P. of three Bayesian methods)

Pearson Correlation Coefficients / Prob > |R| under Ho: Rho=0 / N = 21

	B	G	M
B	1.00000 0.0	-0.29672 0.4274	-0.05205 0.8227
G	-0.29672 0.1915	1.00000 0.0	0.73422 0.0002
M	-0.05205 0.8227	0.73422 0.0002	1.00000 0.0

Table 70
Coefficient Estimates of Various Techniques for Social Characteristics
(GSS Table 9, For CITY)

Variable	GSS[1]	GSS[2]	Geweke[3]	BMA[4]
Log (initial population)	-0.043	-0.009	-0.016	-0.293
Initial per capital income	-0.159	-0.740	0.073	-0.012
Initial % nonwhite	-0.048	-0.006	-0.001	-0.044
Initial unemployment rate	-0.039	-0.175	0.000	0.118
Initial manufacturing share	-0.429	-0.001	-0.002	0.016
Initial median years of schooling	0.060	0.169	0.064	-0.000
Initial segregation index	-0.006	-0.319	0.081	-0.101
Initial segregation * initial % nonwhite	0.529	0.003	-0.009	-0.174
South	-0.296	-0.208	-0.087	-0.059
Central	-0.482	-0.486	-0.311	-0.006
NEast	-0.478	-0.378	-0.214	-0.000
Initial aging population			-0.053	-0.002

1. The GSS's original analysis, the initial year is 1960
2. To be consistent with the data set used in this dissertation, which the initial year is 1970
3. The posterior mean of variables, the initial year is 1970
4. The initial year is 1970

d

VITA

Jui-Chu Tien

Candidate for the Degree of

Doctor of Philosophy

Thesis: BAYESIAN VARIABLE AND MODEL SELECTION IN ECONOMICS: AN APPLICATION IN URBAN ECONOMICS

Major Field: Economics

Biographical:

Education: Graduated from Taichung First High School, Taichung, Taiwan in June 1984; received Bachelor of Business Administration from Tunghai University, Taichung, Taiwan in June 1988; received Master of Business Administration from Louisiana Tech University, Ruston, Louisiana in August 1993. Completed the requirements for the Doctor of Philosophy degree with a major in Economics at Oklahoma State University in (December, 2001).

Experience: Employed as sergeant by Republic of China Army; employed as Mathematical teacher by Ho-Mei Junior High School; employed by Louisiana Tech University, Department of Management Sciences as a graduate teaching assistant; employed by Oklahoma State University, Department of Economics and Legal Studies as a graduate teaching assistant, 1996 to present.

Professional Memberships: American Economics Association, American Statistical Association.