Hierarchical Bayes prediction for the 2008 US Presidential election

Sinha, Pankaj and Bansal, Ashok
Faculty of Management Studies, University of Delhi, Delhi, University of Delhi

13. August 2008

Hierarchical Bayes Prediction for the 2008 US Presidential Election

Pankaj Sinha[1] and Ashok K. Bansal[2]

Abstract

In this paper a procedure is developed to derive the predictive density function of a future observation for prediction in a multiple regression model under hierarchical priors for the vector parameter. The derived predictive density function is applied for prediction in a multiple regression model given in Fair (2002) to study the effect of fluctuations in economic variables on voting behavior in U.S. presidential election. Numerical illustrations suggest that the predictive performance of Fair's model is good under hierarchical Bayes setup, except for the 1992 election. Fair's model under hierarchical Bayes setup indicates that the forthcoming 2008 US presidential election is likely to be a very close election slightly tilted towards Republicans. It is likely that republicans will get 50.90% vote with probability for win 0.550 in 2008 US Presidential Election.

1. Introduction

Consider a prediction problem where the outcomes $x_1, x_2, ..., x_n$ of informative experiments are independent with probability density function $f(x_i \mid \theta_i)$, $i = 1,2,...,n$. The outcome x_{n+1} of a future independent experiment has p.d.f. $f(x_{n+1} \mid \theta_{n+1})$, the parameter θ_{n+1} has same parameter space Θ as that of θ_i ($i = 1,2,..., n$). Our objective is to derive the predictive density function of x_{n+1}, given the outcomes $x_1, x_2, ..., x_n$ of informative experiments for prediction in a multiple regression model. One approach to deal with this prediction problem is to employ hierarchical priors in a Bayesian framework. Hierarchical priors are used when the parameter $\underset{\sim}{\theta}$ is a vector ($\theta_1, \theta_2, ..., \theta_n$) and it is assumed that θ_i ($i = 1,2,..., n$) are distributed independently with common prior distribution $g(\theta_i \mid \lambda)$ and a second stage prior distribution $g(\lambda)$ is placed on it, i.e., on λ.

A hierarchical Bayesian regression model has been found useful in the area of applied econometrics and statistics. Lindley & Smith (1972) initially developed the general Bayesian linear model, which is also known as (linear) hierarchical model. Polasek (1984) developed an empirical Bayes estimation of a 2-stage hierarchical model. Polasek & Potzelberger (1988) carried out robust Bayesian analysis with a hierarchical time series model using Austrian economic data. Berger and Berliner (1986) used $\varepsilon-$contaminated class of priors to represent the uncertainty both in $g(\theta_i \mid \lambda)$ and $g(\lambda)$ to investigate the robustness with respect to hierarchical priors. Aitchison & Dunsmore (1975) illustrates the wide applicability of Bayes predictive approach.

[1] Faculty of Management Studies, University of Delhi, Delhi. (Email: pankaj-sinha@fms.edu)
[2] Department of Statistics, University of Delhi, Delhi

In section 2, we demonstrate the standard Bayesian method to find the predictive density function of a future observation x_{n+1}, given the outcomes of an informative experiment, under hierarchical priors. In section 3, the derived predictive density function is modified for the purpose of prediction in a multiple regression model, assuming that θ_i's are independent and their prior distributions are described in two stages. The expressions for one period forward forecast and predictive interval are obtained in sections 4 and 5.

In section 6, to demonstrate the hierarchical Bayes approach to forecast the 2008 US presidential election, the derived results are applied to the multiple regression model and data given in Fair (2002) for studying the effect of fluctuations in economic variables on voting behavior in U.S. presidential election. Fair (1978) examined the economic determinants of the presidential popular vote. Fair's model has contributed significantly to research into presidential election. The more recent works in the area are found in Berry and Harpham (1996), Erikson and Wlezien (1996), Hibbs (2000) and Fair (2004). Gleisner (1992, 2005) critically examines the Fair's model.

We denote density function $g(.)$ on parameter space Θ (i.e., prior as well as posterior), density function $f(.)$ on the sample observations and $p(.)$ as predictive density function to simplify the notations.

2. Prediction Under Hierarchical Priors

Let $x_1, x_2, ..., x_n$ be independent observations from $f(x_i \mid \theta_i)$, $i = 1,2,..., n$, where θ_i's are independent and their prior distribution may be described in two stages.

Stage1: θ_i 's are conditionally independently distributed as $g(\theta_i \mid \lambda)$ with a common parameter $\lambda \in \Lambda$.

Stage 2: The parameter λ at stage 1 has a proper prior distribution $g(\lambda)$.

Let the future observation x_{n+1} be distributed as $f(x_{n+1} \mid \theta_{n+1})$ and θ_{n+1} has the same parameter space Θ as that of θ_i ($i = 1,2,..., n$).

The predictive density function of the future observation x_{n+1}, given $\underline{x} = \{x_1, x_2, ..., x_n\}$, may be obtained as follows:

$$p(x_{n+1} \mid \underline{x}) = \int_{\Theta} p(x_{n+1} \mid \underline{\theta}) g(\underline{\theta} \mid \underline{x}) d\underline{\theta} \tag{2.1}$$

where,

$$P(x_{n+1} \mid \underline{\theta}) = \int_{\Theta} f(x_{n+1} \mid \theta_{n+1}) g(\theta_{n+1} \mid \underline{\theta}) d\theta_{n+1} \tag{2.2}$$

$$g(\theta_{n+1} \mid \underline{\theta}) = \int_{\Lambda} g(\theta_{n+1} \mid \lambda) g(\lambda \mid \underline{\theta}) d\lambda \tag{2.3}$$

$$g(\lambda \mid \underline{\theta}) = \frac{g(\lambda) \prod_{i=1}^{n} g(\theta_i \mid \lambda)}{\int_{\Lambda} g(\lambda) \prod_{i=1}^{n} g(\theta_i \mid \lambda) d\lambda} \tag{2.4}$$

$$g(\underline{\theta} \mid \underline{x}) = \int_{\Lambda} g(\underline{\theta} \mid \underline{x}, \lambda) g(\lambda) d\lambda \tag{2.5}$$

and

$$g(\underline{\theta} \mid \underline{x}, \lambda) = \frac{f(\underline{x} \mid \underline{\theta}) g(\underline{\theta} \mid \lambda)}{\int_{\Theta} f(\underline{x} \mid \underline{\theta}) g(\underline{\theta} \mid \lambda) d\underline{\theta}} \tag{2.6}$$

$$= \prod_{i=1}^{n} \frac{f(x_i \mid \theta_i) g(\theta_i \mid \lambda)}{\int_{\Theta} \prod_{i=1}^{n} f(x_i \mid \theta_i) g(\theta_i \mid \lambda) d\theta_i}, \tag{2.7}$$

since $x_1, x_2, ..., x_n$ are independent random variables and $\theta_1, \theta_2, ..., \theta_n$ are also assumed to be independent.

Example 2.1

Let $x_1, x_2, ..., x_n$ be independent observations from $N(\theta_i, r), i = 1,2,......,..n$, with mean θ_i and known common precision r. Let the future observation $x_{n+1} \sim N(\theta_{n+1}, r)$. Assume that θ_i 's are independent and their prior distributions are described in two stages (c.f. Berger (1985).

Stage 1: θ_i 's are independent and normally distributed each with mean μ and known precision τ. We have

$$g(\theta_i \mid \mu) = \sqrt{\frac{\tau}{2\pi}} \exp[-\frac{\tau}{2}(\theta_i - \mu)^2] \tag{2.8}$$

Stage 2: the common parameter μ at stage 1 has a normal prior distribution with mean a and precision b; it is represented by

$$g(\mu) = \sqrt{\frac{b}{2\pi}} \exp[-\frac{b}{2}(\mu - a)^2] \tag{2.9}$$

Though the MCMC methods freed the analysts from using conjugate prior distributions for mathematical convenience, the advantage of conjugate prior is that it treats the prior information as if it were a previous sample of the same process.

Let us use the fact that the sample mean provides the sufficient statistic for the unknown mean of the normal population.

Let

$$\bar{\theta} = \frac{1}{n}\sum_{i=1}^{n}\theta_i \quad \text{and} \quad \bar{x} = \frac{1}{n}\sum_{i=1}^{n}x_i ,$$

we find

$$g(\mu \mid \bar{\theta}) = \frac{g(\mu)g(\bar{\theta} \mid \mu)}{\int_{-\infty}^{\infty} g(\mu)g(\bar{\theta} \mid \mu)d\mu}$$

$$= \sqrt{\frac{\tau'}{2\pi}}\exp[-\frac{\tau}{2}(\mu - c)^2] \tag{2.10}$$

$$g(\theta_{n+1} \mid \bar{\theta}) = \int_{-\infty}^{\infty} g(\theta_{n+1} \mid \mu)g(\mu \mid \bar{\theta})d\mu$$

$$= \sqrt{\frac{\tau'_1}{2\pi}}\exp[-\frac{\tau'_1}{2}(\theta_{n+1} - c)^2] \tag{2.11}$$

$$p(x_{n+1} \mid \bar{\theta}) = \int_{-\infty}^{\infty} f(x_{n+1} \mid \theta_{n+1})g(\theta_{n+1} \mid \bar{\theta})d\theta_{n+1}$$

$$= \sqrt{\frac{\tau''}{2\pi}}\exp[-\frac{\tau''}{2}(x_{n+1} - c)^2] \tag{2.12}$$

$$g(\bar{\theta} \mid \bar{x}, \mu) = \frac{f(\bar{x} \mid \bar{\theta})g(\bar{\theta} \mid \mu)}{\int_{-\infty}^{\infty} f(\bar{x} \mid \bar{\theta})g(\bar{\theta} \mid \mu)d\bar{\theta}}$$

$$= \sqrt{\frac{(r+\tau)n}{2\pi}}\exp[-\frac{(r+\tau)n}{2}(\bar{\theta} - \mu_1')^2] \tag{2.13}$$

$$g(\bar{\theta} \mid \bar{x}) = \int_{-\infty}^{\infty} g(\bar{\theta} \mid \bar{x}, \mu)g(\mu)d\mu$$

$$= \sqrt{\frac{l_2}{2\pi}} \exp[-\frac{l_2}{2}(\overline{\theta} - g_1)^2] \qquad (2.14)$$

Thus the predictive density function of a future observation x_{n+1}, given $\overline{x}$, is given by

$$p(x_{n+1} \mid \overline{x}) = \int_{-\infty}^{\infty} p(x_{n+1} \mid \overline{\theta}) \xi(\overline{\theta} \mid \overline{x}) d\overline{\theta}$$

$$= \sqrt{\frac{l_4}{2\pi}} \exp[-\frac{l_4}{2}(x_{n+1} - g_2)^2]$$

(2.15)

Where,

$$l_4 = \frac{\tau'' l_2}{l_3 + l_2}, \quad l_3 = \tau''(\frac{n\tau}{\tau'})^2, \quad l_2 = \frac{(r+\tau)bn}{l_1 + b}, \quad l_1 = \frac{n\tau^2}{r+\tau}, \quad g_1 = \frac{r\overline{x} + \tau a}{r+\tau},$$

$$g_2 = (\frac{ba + n\tau g_1)}{\tau'}, \quad \tau'' = \frac{r\tau'_1}{r + \tau'_1}, \quad \tau'_1 = \frac{\tau\tau'}{\tau_1}, \quad \tau_1 = \tau + \tau', \quad \tau' = n\tau + b,$$

and $c = (n\tau\overline{\theta} + ba)/(n\tau + b)$

3. Prediction in the Regression Model

Let the informative experiments assume normal regression of endogenous variable x on $m-1$ exogenous variables $t_{2,} t_3 ......t_m$.

$$x_i = \beta_1 + \beta_2 t_{2i} + ... + \beta_m t_{mi} + \varepsilon_i, \; i = 1,2,..., n \qquad (3.1)$$

where, each $\varepsilon_i \sim N(0, \; \sigma^2)$ with mean 0 and variance σ^2 so that

$$E(x_i) = T_i \underline{\beta}$$

with $\underline{\beta} = [\, \beta_1 \; \beta_2 ....\beta_m]'$ and $T_i = [1 \;\; t_{2i} \;\; t_{3i} ......t_{mi}]$.

The informative experiments yield observations $x_1, x_2, ..., x_n$, which are independently distributed having normal p.d.f. with respective means $\theta_1, \theta_2, ..., \theta_n$ and common variance σ^2. Here $\theta_i (= T_i \; \underline{\beta})$.

Consider the data set represented by

$$T = \begin{bmatrix} 1 & t_{21} & t_{31} \ldots\ldots\ldots\ldots t_{m1} \\ 1 & t_{22} & t_{32} \ldots\ldots\ldots\ldots t_{m2} \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ 1 & t_{2n} & t_{3n}^2 \ldots\ldots\ldots\ldots t_{mn} \end{bmatrix}, \qquad X = \begin{bmatrix} x_1 \\ x_2 \\ .. \\ x_n \end{bmatrix}.$$

The least square estimate of $\underline{\beta}$ is given by $\underline{\hat{\beta}} = (T'T)^{-1}T'X$.

$\underline{\hat{\beta}}$ has a multivariate normal distribution, i.e. $\underline{\hat{\beta}} \sim N_m(\underline{\beta},\ \sigma^2(T'T)^{-1})$ and $T_i\hat{\beta}$ has a normal distribution, i.e., $T_i\,\underline{\hat{\beta}} \sim N(T_i\,\underline{\beta},\ \sigma^2 T_i(T'T)^{-1}T_i')$.

Thus $\frac{1}{n}\sum_{i=1}^{n} T_i\hat{\beta} \sim N(\frac{1}{n}\sum_{i=1}^{n} T_i\beta,\ \frac{\sigma^2}{n^2}\sum_{i=1}^{n} T_i(T'T)^{-1}T_i')$.

Note that $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}\sum_{i=1}^{n} T_i\hat{\beta}$ is a sufficient statistic for $\overline{\theta}$ $(=\frac{1}{n}\sum_{i=1}^{n}\theta_i)$, where $\theta_i = T_i\beta$.

We have $\overline{x} \sim N(\overline{\theta}, \overline{p}\frac{\sigma^2}{n})$ with mean $\overline{\theta}$ and variance $\overline{p}\frac{\sigma^2}{n}$, where $\overline{p} = \frac{1}{n}\sum_{i=1}^{n} T_i(T'T)^{-1}T_i'$

The precision of $\overline{x}$ is given by

$$\frac{1}{V(\overline{x})} = \frac{n}{\overline{p}\sigma^2} = kr \ \text{ where, } k = \frac{n}{\overline{p}} \text{ and } r = \frac{1}{\sigma^2} \ .$$

Thus $\overline{x} \sim N(\overline{\theta}, kr)$ with mean $\overline{\theta}$ and precision kr.

Let the outcome x_{n+1} of future experiment be also identically distributed with mean $\theta_{n+1}(=T_{n+1}\,\underline{\beta})$ and precision $k_1 r$, i.e.,

$$x_{n+1} = T_{n+1}\hat{\beta} \sim N(\theta_{n+1},\ k_1 r), \text{ where } k_1 = (T_{n+1}(T'T)^{-1}T_{n+1}')^{-1}.$$

Therefore, the predictive density function of future observation x_{n+1}, when the hierarchical prior distribution for $\theta_i\,(=T_i\,\underline{\beta})$ is given by equations (2.8) and (2.9) and if $r=\frac{1}{\sigma^2}$ is assumed to be known, may be easily rewritten as

$$p(x_{n+1}\mid\bar{x})=\sqrt{\frac{l_4}{2\pi}}\exp\left[-\frac{l_4}{2}(x_{n+1}-g_2)^2\right] \tag{3.2}$$

Where,

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}T_i\hat{\beta}\,,\quad k_1=(T_{n+1}(T'T)^{-1}T'_{n+1})^{-1},\quad k=\frac{n}{\bar{p}},\quad \bar{p}=\frac{1}{n}\sum_{i=1}^{n}T_i(T'T)^{-1}T_i{}'$$

$$l_4=\frac{\tau''l_2}{l_3+l_2},\quad l_3=\tau''\left(\frac{n\tau}{\tau'}\right)^2,\quad l_2=\frac{(kr+n\tau)b}{l_1+b},\quad l_1=\frac{n^2\tau^2}{kr+n\tau}$$

$$g_2=\left(\frac{ba+n\tau g_1)}{\tau'}\right.,\quad g_1=\frac{kr\bar{x}+n\tau a}{kr+n\tau},\quad \tau''=\frac{k_1r\tau'_1}{k_1r+\tau'_1},\quad \tau'_1=\frac{\tau\tau'}{\tau_1},\quad \tau_1=\tau+\tau',$$

$\tau'=k\tau+b$.

4. One -Period Forward Forecast

On the basis of observations $x_1,x_2,...,x_n$, the one -period forward forecast can be expressed as

$$\hat{X}_n(1)=E[x_{n+1}\mid x_n,x_{n-1},.......x_1]$$

$$=\int_{-\infty}^{\infty}x_{n+1}p(x_{n+1}\mid\bar{x})dx_{n+1}$$

$$=g_2 \tag{4.1}$$

where,

$$g_2=\left(\frac{ba+k\tau g_1)}{\tau'}\right.,\quad g_1=\frac{kr\bar{x}+n\tau a}{kr+n\tau},\quad \bar{x}=\frac{1}{n}\sum_{i=1}^{n}T_i\hat{\beta},\ \tau'=k\tau+b,\ k=\frac{n}{\bar{p}},$$

$\hat{\beta}$ is the least squares estimate of β and $\bar{p}=\frac{1}{n}\sum_{i=1}^{n}T_i(T'T)^{-1}T_i{}'$.

5. Predictive Interval

Let us denote

$$\phi(x_{n+1}) = \frac{1}{\sqrt{2\pi}} \exp[-\frac{x_{n+1}^2}{2}] \quad \text{and} \quad \Phi(q) = \int_{-\infty}^{q} \phi(x_{n+1}) dx_{n+1} \ . \tag{5.1}$$

Then the probability $P[x_{n+1} > q \mid \bar{x}]$ is given by

$$P[X_{n+1} > q \mid \bar{x}] = \int_{q}^{\infty} p(x_{n+1} \mid \bar{x}) dx_{n+1} = [1 - \Phi(q^*)] \quad ,$$

$$\text{where ,} \quad q^* = \sqrt{l_4}\,(q - g_2) \ . \tag{5.2}$$

6. Illustration

Consider the following modified model given by Fair (2002) for studying the influence of fluctuations in economic variables on voting behavior in U.S. presidential election.

$$E(Vote) = \beta_1 + \beta_2 Party + \beta_3 Duration + \beta_4 Person + \beta_5 War + \beta_6 Growth + \beta_7 Inflation + \beta_8 Goodnews \tag{6.1}$$

The notation for the above regression equation is as follows:

Vote = Incumbent share of two party vote. Incumbent vote is divided by the Democratic plus Republican vote

Party = 1 if there is a Democratic incumbent at the time of election and –1 if there is a Republican incumbent

Duration = 0 if the incumbent party has been in power for one term, 1 if the incumbent party has been in power for two consecutive terms, 1.25 for three consecutive terms, 1.50 for four consecutive terms, and so on.

Person = 1 if incumbent is running for election and 0 otherwise.

War = 1 for the elections of 1920, 1944 and 1948, and 0 otherwise.

Growth = growth rate of real per capita GDP in the first three quarters of the election year (annual rate)

Inflation = absolute value of the growth rate of the GDP deflator in the first 15 quarters of the administration (annual rate) except for 1920, 1944, 1948, where the values are zero.

Goodnews = number of quarters in the first 15 quarters of the administration in which the growth rate of real per capita GDP is greater than 3.2 percent at an annual rate except for 1920, 1944, and 1948, where the values are zero.

Table 6.4 gives Fair' s data on quadrennial presidential elections in the United States from 1916 to 2004. Quarterly data on nominal GDP, real GDP and population are used to construct the variables Growth, Inflation and Goodnews. The economic data and formulation for construction of data on the variables are explained in Fair (2002, 2004).

Let us denote the variable Vote by x, and variables Party, Duration, Person. War, Growth, Inflation and Goodnews by $t_2, t_3, t_4, t_5, t_6, t_7$ and t_8, respectively. Since each election year is unique and its result is independent of its previous and next election results, the equation (6.1) can be written in the form of equation (3.1) and the results derived in equations (3.2), (4.1) and (5.2) can be easily applied for obtaining predictive density function, one period forward forecast and probability for win $P[x_{n+1} > 50.0 \mid \bar{x}]$. We recursively estimate the model and evaluate the out-of-sample one period ahead probability forecast.

The parameters $\beta = (\beta_1, \beta_2, \dots \beta_8)'$ of the model are estimated by the least squares method from the data set given in Table 6.4.These estimates are summarized in Table 6.3.

The precision $r = \dfrac{1}{\sigma^2}$ is assumed to be known, we take $r = \dfrac{1}{\hat{\sigma}^2} = \dfrac{n-8}{RSS}$ as a true value, where RSS = $(X - T\hat{\beta})'(X - T\hat{\beta})$. The estimates of parameters of prior distribution are made on the basis of results of the informative experiments. We take the first stage prior for the unknown mean $\bar{\theta}$ as $N(\mu, \tau)$, where $\tau = \dfrac{n-1}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$ and n is the number of sample observations.

The second stage prior on μ is distributed as $N(a, b)$ with mean a and precision b. Setting $a = T_{n+1}\hat{\beta}$ and $b = r\,(T_{n+1}(T'T)^{-1}T'_{n+1})^{-1}$, the one period forward forecast values, prediction errors and $P[x_{n+1} > 50.0 \mid \bar{x}]$ are summarized in Tables 6.0, 6.1 and 6.2. We find that the predictive performance of the model is very good with the above values of the parameters.

For the sample period 1916 –2000 ($n = 22$), the root mean square error of one period forward forecast is 3.18 and the Theil inequality coefficient is near zero (0.00114). The Theil inequality coefficient for all other sample periods (1916-1996 to 1996-1880) is also near zero. Root mean square error of one period forward forecast is 3.196 and 3.384 for the sample periods

1916-2000 and 1916-1996, respectively. It is below 2.1 for all other sample periods. This suggests the predictive performance of the model is good.

For the 2000 election using sample observations 1916-1996, the model predicted victory for Democratic Party candidate Mr. Al Gore by a narrow margin (50.948) with probability 0.552. For the 2004 election using sample observations 1916-2000, it predicted victory for President Bush by a fairly comfortable margin (54.463) with probability 0.736. Though President Bush won both the elections, the margin in 2000 election was narrow (50.265). The model prediction was good for the 1996 election when it predicted victory for President Clinton (52.633) with probability 0.646 using sample observations 1916-1992, President Clinton could secure 54.736 percentage of vote share. The model predictions are also true for the 1988, 1984 and 1980 elections. The model predicted victory for the incumbent in the 1988 and 1984 elections, with one period forward forecasts 51.836 (probability to win 0.596) and 60.223 (probability to win 0.991), respectively. Using sample observations 1916-1976, the model predicted defeat of the incumbent in the election of 1980 with one period forward forecast 48.672 and probability for victory 0.431.

The 1992 election is the most problematic election for the model. It predicted victory for President Bush (54.042) with a probability 0.707 but he lost to Mr. Clinton by a fairly large margin (46.545). Fair (1996) tries to explain this error in prediction.

2008 US presidential election

Table 6.2 gives the hierarchical Bayes forecast on Fair's vote model for the 2008 election. It suggests that the 2008 presidential election is likely to be a close election slightly tilted towards the republicans if the GDP, inflation and Goodnews remain at the current level (July 2008) of 1.0%, 3.0% and 3 respectively. At this level of GDP and inflation, it is likely that republicans will get 50.90% vote with probability for win 0.550.

Reference

Aitchison, J. and Dunsmore, I. R. (1975) *Statistical Prediction Analysis,* Cambridge, Cambridge University Press.

Berger, J. O. (1985) *Statistical Decision Theory and Bayesian Analysis* Springer, New York.

Berger, J. O. & Berliner, M. (1986): Robust Bayes and empirical Bayes analysis with $\in$-contaminated priors, *The Annals of Statistics,* 14 (2), pp. 461-486.

Berry, B., Elliot, E., and Harpham, E. J. (1996) The yield curve as an electoral bellwether, *Technical forecasting and social change,* 51, pp. 281-294.

Erikson, R. S., and Wlezien, C. (1996) Of time and presidential election forecasts PS: *Political Science and politics,* 31, pp. 37-39.

Fair, R. C. (1978) The effect of economic events on votes for president, *Review of Economics and Statistics,* 60, pp. 159-173.

Fair, R. C. (1996) The effect of economic events on votes for president: 1992 update, *Political Behavior,* 18, pp. 119-139

Fair, R. C. (2002) *Predicting Presidential Elections and Other Things,* Stanford: Stanford University Press.

Fair, R. C. (2004) A vote equation and the2004 election, Website: http:// fairmodel.econ.yale.edu/vote2004

Gleisner, R. F. (1992) Economic developments of presidential election: The Fair model, *Political Behavior*, 14, pp. 383-394.

Gleisner, R. F. (2005) Comments for presentation at the roundtable on Fair's presidential vote equation in International Symposium on forecasting, San Antonio, June 14, 2005

Hastings, C. (1955) *Approximation for Digital Computers,* Princeton, NJ, Princeton University Press.

Hibbs, D. A. (2000) Bread and peace voting in U.S. presidential election, *Public Choice,* 104, pp. 149-180.

Lindley, D. V. and Smith, A. F. M.(1972) Bayes estimates for the linear model, (with discussion). *Journal of Royal Statistical Society,* B 34, pp. 1-41.

Polasek, W. (1984) Multivariate regression systems: Estimation and sensitivity analysis for two - dimensional data. *Robustness in Bayesian Statistics,* (J. Kadane, ed.) Amsterdam: North-Holland, pp. 1-41.

Polasek, W. and Potzelberger, K. (1988) Robust Bayesian Analysis in Hierarchical Models, *Bayesian Statistics 3*, Oxford University Press, pp. 377-394

Table –6.0
One Period Forward Hierarchical Bayes Forecast Estimates for Fair's Vote model

Year	**Sample** n	**Forecast Vote share of Incumbent %**	**Actual Vote share of Incumbent %**	**Forecast Error**	$r.m.s.$ **Error**	**Prob. for win*** $P\ [\ x_{n+1} > 50 \mid \bar{x}]$
2004	**22** (1916-2000)	**54.059**	**51.233**	2.826	3.187	**0.716**
2000	**21** (1916-1996)	**50.948**	**49.735**	1.213	3.271	**0.552**
1996	**20** (1916-1992)	**52.577**	**54.736**	2.159	3.385	**0.641**
1992	19 (1916-1988)	**53.738**	**46.545**	-7.193	2.093	**0.692**
1988	**18** (1916-1984)	**51.836**	**53.902**	2.066	1.953	**0.596**
1984	**17** (1916-1980)	**60.228**	**59.17**	-1.057	1.370	**0.911**
1980	**16** (1916-1976)	**48.672**	**44.697**	3.975	1.321	**0.431**

Table –6.1
One Period Forward Hierarchical Bayes Forecast Estimates for Fair's Vote model

Year	Sample n	Prior Parameters a	b	τ	$r = \frac{1}{\hat{\sigma}^2}$	**Forecast** Vote share of Incumbent %	**Actual** Vote share of Incumb-ent %	**Thiel Inequality Coeff.**	**Prob. for win** $P[x_{n+1} \mid \bar{x} > 50.0]$
2004	**22** (1916-2000)	58.269	0.427	0.021	0.178	**54.059**	**51.233**	0.00114	**0.716**
2000	**21** (1916-1996)	48.58	0.643	0.0201	0.171	**50.948**	**49.735**	0.00117	**0.552**
1996	**20** (1916-1992)	53.078	0.343	0.0212	0.161	**52.577**	**54.736**	0.00123	**0.641**
1992	19 (1916-1988)	55.353	0.525	0.0188	0.370	**53.738**	**46.545**	0.00074	**0.692**
1988	**18** (1916-1984)	51.428	1.342	0.0178	0.400	**51.836**	**53.902**	0.00074	**0.596**
1984	**17** (1916-1980)	63.343	1.606	0.0176	0.697	**60.228**	**59.17**	0.00049	**0.911**
1980	**16** (1916-1976)	45.685	0.745	0.0177	0.648	**48.672**	**44.697**	0.00046	**0.431**

Table- 6.2
Hierarchical Bayes Forecast on Fair's Vote Model for the 2008 Election

Sample 1916 – 2004 **Number of observations $n = 23$**

Year	Growth	Inflation	Goodnews	Prior Parameters			r	Forecast Vote share of Incumbent %	Probability for Win
				a	b	τ			
April 2007	1.9	3.3	1	46.808	0.428	0.02193	0.155	50.78	**0.543**
July 2008	1.0	3.0	3	48.543	0.682	0.02193	0.155	50.90	**0.550**

Table –6.3

Least Squares Estimates of Fair's Vote Model

Election Year	**Sample** n	constant	Party	Duration	Person	War	Growth	Inflati-on	Good News
		$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	$\hat{\beta}_7$	$\hat{\beta}_8$
2008	**23** (1916-2004)	47.264	-2.676	-3.330	3.296	5.614	0.680	-0.657	1.075
2004	**22** (1916-2000)	49.608	-2.713	-3.628	3.251	3.855	0.691	-0.775	0.837
2000	**21** (1916-1996)	49.405	-2.808	-3.641	3.451	4.043	0.697	-0.763	0.827
1996	**20** (1916-1992)	48.594	-2.914	-3.420	3.441	4.699	0.703	-0.714	0.896
1992	19 (1916-1988)	49.543	-3.251	-2.104	5.319	1.238	0.738	-0.866	0.558
1988	**18** (1916-1984)	48.843	-3.139	-2.164	5.520	1.754	0.728	-0.837	0.619
1984	**17** (1916-1980)	47.616	-3.463	-2.255	5.618	3.409	0.768	-0.708	0.764
1980	**16** (1916-1976)	47.645	-3.354	-2.357	5.585	3.412	0.762	-0.662	0.759

TABLE- 6.4

Fair (2002) Data on U.S. Presidential Elections, 1916-2000

Year	Vote	Party	Duration	Person	War	Growth	Inflation	Good news
1916	51.682	1	0.00	1	0	2.229	4.252	3
1920	36.119	1	1.00	0	1	-11.463	0.000	0
1924	58.244	-1	0.00	1	0	-3.872	5.161	10
1928	58.820	-1	1.00	0	0	4.623	0.183	7
1932	40.841	-1	1.25	1	0	-14.557	7.160	4
1936	62.458	1	0.00	1	0	11.677	2.454	9
1940	54.999	1	1.00	1	0	3.611	0.055	8
1944	53.774	1	1.25	1	1	4.433	0.000	0
1948	52.370	1	1.50	1	1	2.858	0.000	0
1952	44.595	1	1.75	0	0	0.840	2.316	6
1956	57.764	-1	0.00	1	0	-1.394	1.930	5
1960	49.913	-1	1.00	0	0	0.417	1.963	5
1964	61.344	1	0.00	1	0	5.109	1.267	10
1968	49.596	1	1.00	0	0	5.070	3.156	7
1972	61.789	-1	0.00	1	0	6.125	4.813	4
1976	48.948	-1	1.00	0	0	4.026	7.579	4
1980	44.697	1	0.00	1	0	-3.594	7.926	5
1984	59.170	-1	0.00	1	0	5.568	5.286	8
1988	53.902	-1	1.00	0	0	2.261	3.001	4
1992	46.545	-1	1.25	1	0	2.223	3.333	2
1996	54.736	1	0.00	1	0	2.712	2.146	4
2000	50.265	1	1.00	0	0	1.603	1.679	7

TABLE- 6.5
Fair (2007) Revised Data on U.S. Presidential Elections, 1916-2004

Year	Vote	Party	Duration	Person	War	Growth	Inflation	Good news
1916	51.682	1	0.00	1	0	2.229	4.252	3
1920	36.119	1	1.00	0	1	-11.463	0.000	0
1924	58.244	-1	0.00	1	0	-3.872	5.161	10
1928	58.820	-1	1.00	0	0	4.623	0.183	7
1932	40.841	-1	1.25	1	0	-14.499	7.200	4
1936	62.458	1	0.00	1	0	11.765	2.497	9
1940	54.999	1	1.00	1	0	3.902	0.081	8
1944	53.774	1	1.25	1	1	4.279	0.000	0
1948	52.370	1	1.50	1	1	2.579	0.000	0
1952	44.595	1	1.75	0	0	0.691	2.362	7
1956	57.764	-1	0.00	1	0	-1.451	1.935	5
1960	49.913	-1	1.00	0	0	0.377	1.967	5
1964	61.344	1	0.00	1	0	5.109	1.260	10
1968	49.596	1	1.00	0	0	5.043	3.139	7
1972	61.789	-1	0.00	1	0	5.914	4.815	4
1976	48.948	-1	1.00	0	0	3.751	7.630	5
1980	44.697	1	0.00	1	0	-3.597	7.831	5
1984	59.170	-1	0.00	1	0	5.440	5.259	8
1988	53.902	-1	1.00	0	0	2.178	2.906	4
1992	46.545	-1	1.25	1	0	2.662	3.280	2
1996	54.736	1	0.00	1	0	3.121	2.062	4
2000	50.265	1	1.00	0	0	1.219	1.605	8
2004	51.233	-1	0.0	1	0	2.690	2.325	1
Jan2007	--	-1	1.0	0	0	1.7	3.4	1
April 2007	---	-1	1.0	0	0	1.9	3.3	1
July 2008		-1	1.0	0	0	1.0	3.0	3