Analysis of the Australian Election Study Using Bayesian Quantile Regression Models

Charles Au *　　S. T. Boris Choy †

Abstract
Quantile regression is useful for modeling the conditional quantile of the response variable. Recently, quantile regression has also been applied to discrete choice models, where the response variable is binary or ordinal. They can be estimated using the Bayesian Markov chain Monte Carlo (MCMC) approach when the error terms are assumed to follow, for example, the asymmetric Laplace distribution. This paper proposes the application of Bayesian quantile regression models to survey data from the Australian Election Study (AES). The binary and ordinal quantile regression models will be used for investigating the factors that influence Australian voters' choice for certain political parties and their level of interest in politics generally. In addition, to assist with the interpretations of regression coefficients, this paper proposes to calculate the marginal effects of the explanatory variables. The main objectives are to investigate the differences in the coefficients estimates and marginal effects of the regression models at various quantile levels. Comparisons will also be made to binary and ordinal probit models.

Key Words: Markov chain Monte Carlo; quantile regression; discrete choice models; Bayesian inference; Australian Election Study (AES)

1. Introduction

Quantile regression, initially proposed by Koenker and Bassett (1978), is often used as an alternative to ordinary least squares regression. With the increased popularity of the Bayesian approaches in the past three decades, Bayesian Markov chain Monte Carlo (MCMC) methods have been extended to quantile regression, where the error term is assumed to follow the asymmetric Laplace distribution (Yu and Moyeed, 2001). This provided an alternative method for statistical inference to a non-parametric problem. Kozumi and Kobayashi (2011) proposed using the normal-exponential mixture representation of the asymmetric Laplace distribution to facilitate the use of the Gibbs sampler in estimating Bayesian quantile regression models. As an alternative to the asymmetric Laplace distribution, Wichitaksorn, Choy, and Gerlach (2014) proposed another class of skew distributions that could also be used for quantile regression models.

Until recently, quantile regression models have been applied to situations where the response variable is continuous. In Albert and Chib (1993), Bayesian methods had been applied to regression models where the response variable is binary or ordered data. Benoit and van den Poel (2012) and Rahman (2016) considered the use of Bayesian binary and ordered quantile models respectively, which looked at the conditional quantiles of the response variable rather then the conditional mean. Bayesian binary quantile models have been applied to areas such as education (Mollica and Petrella, 2016) and environmental studies (Lavín, Flores, and Ibarnegaray, 2017).

*Discipline of Business Analytics, University of Sydney, Australia.
†Discipline of Business Analytics, University of Sydney, Australia

However, one difficulty of using quantile regression models for binary and ordinal data is that the coefficients are difficult to interpret. Greene (2012) has proposed calculating the marginal effects of variables in the binary and ordinal probit and logistic regression models. The marginal effects of the explanatory variables, rather than the regression coefficients, could be interpreted like the regression coefficients in a linear regression. This means that a unit increase in a continuous explanatory variable (holding all else constant) would lead to a change in the predicted probability of the response variable by the marginal effect for that variable. For a binary explanatory variable, the marginal effect is given by the predicted probability when that variable takes a value of 1 minus the predicted probability when that variable takes a value of 0.

This paper proposes an extension of the method of calculating marginal effects in Greene (2012) to quantile regression models for binary and ordinal response variables. Marginal effects of variables in quantile regression models will be calculated using the MCMC samples for the regression coefficients, which are then converted to MCMC samples for the marginal effects. This paper illustrates the application of this method to the Australian Election Study (AES) 2016 survey data (McAllister et al., 2016). This dataset consists of the responses of 2818 Australians to questions relating to the 2016 Federal election and their political views generally.

The remaining parts of this paper are as follows. Section 2 gives a brief overview of quantile regression of Koenker and Bassett (1978) and its recent Bayesian extensions. Section 3 illustrates the Bayesian binary and ordered quantiles models that will be used in the analysis of the AES dataset. Section 4 shows the application of the models to the AES data. Section 5 concludes this paper.

2. Quantile regression

In this paper, the use of the quantile regression for binary and ordered responses is proposed. The aim is to model the conditional quantiles of the response variable instead of the conditional mean. The quantile at τ $(0 < \tau < 1)$ of a random variable Y (whose realisation is y and whose cumulative distribution function is F) is defined as

$$F^{-1}(\tau) = \inf\{y : F(y) \geq \tau\}. \tag{1}$$

Suppose the response variable $\boldsymbol{y}^*$ ($n \times 1$ vector) is related to the explanatory variables $\mathbf{X}$ ($n \times p$ matrix) at a particular value of τ, such that

$$\boldsymbol{y}^* = \mathbf{X}\boldsymbol{\beta}_\tau + \boldsymbol{\varepsilon}, \tag{2}$$

where $\boldsymbol{\beta}_\tau$ ($p \times 1$ vector) are the regression coefficients for the quantile model at a particular value of τ and $\boldsymbol{\varepsilon}$ ($n \times 1$ vector) are the error terms.

Instead of minimising the sum of squared errors, the loss function in quantile regression of Koenker and Bassett (1978, p. 38) to minimise is

$$\min_{\boldsymbol{\beta}_\tau \in \mathbb{R}^p} \left[\sum_{i \in \{i : y_i \geq \boldsymbol{x}_i^\intercal \boldsymbol{\beta}_\tau\}} \tau |y_i - \boldsymbol{x}_i^\intercal \boldsymbol{\beta}_\tau| + \sum_{i \in \{i : y_i < \boldsymbol{x}_i^\intercal \boldsymbol{\beta}_\tau\}} (1-\tau)|y_i - \boldsymbol{x}_i^\intercal \boldsymbol{\beta}_\tau| \right]. \tag{3}$$

where $\boldsymbol{x}_i$ ($p \times 1$ vector) are the regressors. By defining a check function $\rho_\tau(\varepsilon_i) = \varepsilon_i(\tau - \mathbb{I}(\varepsilon_i < 0))$, where $\mathbb{I}(\cdot)$ is an indicator function, Equation (3) can be rewritten as

$$\min_{\boldsymbol{\beta}_\tau \in \mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau(y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}_\tau). \tag{4}$$

Even though quantile regression models are usually solved using nonparametric method, such models could easily be solved using parametric methods in the Bayesian context, by recognising the fact that minimising the loss function is the same as maximising the likelihood function of the asymmetric Laplace distribution (Yu and Moyeed, 2001). In quantile regression models, the error terms ε_i (for all $i = 1, \ldots, n$) are assumed to follow the asymmetric Laplace distribution of Yu and Zhang (2005). The probability density function of ε_i for a particular value of τ is

$$f_{\varepsilon_i}(\varepsilon_i) = \tau(1-\tau)\exp\left[-\rho_\tau(\varepsilon_i)\right], \tag{5}$$

where $\rho_\tau(\cdot)$ is the check function mentioned above. Since ε_i has location 0, scale 1 and skewness τ, then $\varepsilon_i \sim \text{ALD}(0, 1, \tau)$. It should be noted that while τ represents the skewness parameter in the asymmetric Laplace distribution, it coincides with the parameter for quantile regression that is previously defined.

Kozumi and Kobayashi (2011) showed that this distribution can be written into a hierarchical form as

$$\begin{aligned} \varepsilon_i = g_\tau \lambda_i + h_\tau \sqrt{\lambda_i} u_i, \quad u_i \overset{\text{iid}}{\sim} \mathcal{N}(0,1), \quad \lambda_i \overset{\text{iid}}{\sim} \text{Exp}(1), \\ g_\tau = \frac{1-2\tau}{\tau(1-\tau)}, \quad h_\tau^2 = \frac{2}{\tau(1-\tau)}, \end{aligned} \tag{6}$$

where g_τ and h_τ are fixed constants given τ. Similar to the scale mixtures of normal representation in Andrews and Mallows (1974), the hierarchical form in Equation (6) facilitates the use of the Gibbs sampler in MCMC algorithms.

3. Bayesian quantile regression models for discrete choice data

This section presents the quantile regression models for binary and ordered responses, along with their implementations using Bayesian MCMC algorithms.

3.1 Model for binary responses

Let y_i, $i = 1, \ldots, n$, be the binary response variable taking only the values 1 or 0. The latent variable y_i^* is related to y_i as follows.

$$y_i = \begin{cases} 1, \text{ if} & y_i^* \geq 0, \\ 0, \text{ if} & y_i^* < 0, \end{cases} \tag{7}$$

where the latent variable y_i^* is related to the regressors $\mathbf{x}_i$ as follows

$$y_i^* = \mathbf{x}_i^\top \boldsymbol{\beta}_\tau + \varepsilon_i, \quad \varepsilon_i \overset{\text{iid}}{\sim} \text{ALD}(0, 1, \tau). \tag{8}$$

To ensure identifiability of this model, the scale of the error terms ε_i is fixed, and is equal to 1. Since the error terms follow the asymmetric Laplace distribution, it is possible to use the mixture of normal and exponential variables shown in Equation (6).

To complete the Bayesian paradigm, the following vague prior will be used for $\boldsymbol{\beta}_\tau$.

$$\boldsymbol{\beta}_\tau \sim \mathcal{N}_p(\mathbf{0}, 10000\mathbf{I}_p). \tag{9}$$

The MCMC algorithm for the quantile regression model at a fixed quantile τ for binary responses is given below. The parameters and latent variables are simulated from their full conditional distributions as follows.

- Simulate $\boldsymbol{\beta}$ from a multivariate normal distribution.
- Simulate y_i^* for all $i = 1, \ldots, n$ from a univariate truncated normal distribution.
- Simulate λ_i for all $i = 1, \ldots, n$ from a generalised inverse Gaussian distribution. Since the generalised inverse Gaussian distribution used here is a special case of the inverse Gaussian distribution, the simulation method in Michael, Schucany, and Haas (1976) is used.

For comparison purposes, the binary probit model will also be considered. This means skipping the step for simulating λ_i, as $\lambda_i = 1$ for all $i = 1, \ldots, n$, when estimating the binary probit model.

3.2 Model for ordered responses

Let y_i, $i = 1, \ldots, n$, be the ordered response variable taking values from $j = 1, \ldots, J$, where J is the total number of ordered categories, as follows.

$$y_i = \begin{cases} 1, \text{ if } & \gamma_{\tau,0} < y_i^* \leq \gamma_{\tau,1}, \\ 2, \text{ if } & \gamma_{\tau,1} < y_i^* \leq \gamma_{\tau,2}, \\ \ldots & \\ J, \text{ if } & \gamma_{\tau,J-1} < y_i^* \leq \gamma_{\tau,J}, \end{cases} \tag{10}$$

where $\gamma_{\tau,j}$ are the cut-points, with $\gamma_{\tau,0} = -\infty$ and $\gamma_{\tau,J} = \infty$. For the model to be identifiable, the cut-point $\gamma_{\tau,1} = 0$ is used. To speed up the MCMC algorithm, the following transformation of the cut-points is adopted (see Rahman, 2016),

$$\delta_{\tau,j} = \log(\gamma_{\tau,j-1} - \gamma_{\tau,j}), \quad 2 \leq j \leq J-1. \tag{11}$$

At a fixed quantile level τ, the latent variable y_i^* $(i = 1, \cdots, n)$ is related to the explanatory variables $\mathbf{x}_i$ as follows

$$y_i^* = \mathbf{x}_i^\mathsf{T}\boldsymbol{\beta}_\tau + \varepsilon_i, \quad \varepsilon_i \overset{\text{iid}}{\sim} \text{ALD}(0, 1, \tau). \tag{12}$$

The only cut-points that need to be estimated are $\boldsymbol{\delta}_\tau = (\delta_{\tau,2}, \delta_{\tau,3}, \ldots, \delta_{\tau,J})^\mathsf{T}$. Similar to binary quantile regressions, the scale of the error terms ε_i in ordinal quantile regressions is set to 1 to ensure identifiability.

The hierarchical structure for the error terms ε_i shown in Equation (6) can be used to facilitate the MCMC algorithm.

As with the binary quantile model, to complete the Bayesian paradigm, the following vague prior will be used for $\boldsymbol{\beta}_\tau$.

$$\boldsymbol{\beta}_\tau \sim \mathcal{N}_p(\mathbf{0}, 10000\mathbf{I}_p). \tag{13}$$

The prior chosen for $\boldsymbol{\delta}_\tau$ is

$$\delta_{\tau,j} \sim 1 \cdot \mathbb{I}(\delta_{\tau,j} \leq 3), \quad 2 \leq j \leq J-1. \tag{14}$$

For computational purposes, $\boldsymbol{\delta}_\tau$ will be limited so that its value is smaller or equal to $\log 3$. This is because letting $\boldsymbol{\delta}_\tau$ take unrestricted values would possibly lead to $\boldsymbol{\delta}_\tau$ being very large, and model parameters become difficult to simulate in an MCMC algorithm.

The MCMC algorithm for the quantile regression model at a fixed quantile τ for ordinary responses is given below. The parameters and latent variables are simulated from their full conditional distributions as follows.

- Simulate $\boldsymbol{\beta}_\tau$ from a multivariate normal distribution.
- Simulate y_i^* for all $i = 1, \ldots, n$ from a univariate truncated normal distribution.
- Simulate λ_i for all $i = 1, \ldots, n$ from a generalised inverse Gaussian distribution. As in the case with the binary quantile model, the simulation method in Michael, Schucany, and Haas (1976) is also used here.
- Simulate $\boldsymbol{\delta}_\tau$ from a non-standard distribution.

Since the full conditional distribution for $\boldsymbol{\delta}_\tau$ is non-standard, a random-walk Metropolis-Hastings step is used. This is similar to the MCMC algorithm presented in Rahman (2016), except the cut-points $\boldsymbol{\delta}_\tau$ are restricted to less than or equal to 3.

Once again, for comparison purposes, the ordered probit model is estimated. This model can be implemented by setting $\lambda_i = 1$ for all $i = 1, \ldots, n$, and the step for simulating λ_i is thus omitted.

4. Applications

Two empirical applications of quantile regression to the Australian Election Study (AES) 2016 will be considered. The first application is to use a binary quantile regression model for investigating the impacts of factors contributing the House of Representatives vote for the Australian Greens in the 2016 Federal election. The second application is an ordinal quantile regression model for the level of voters' interest in politics generally.

There are 2818 responses on AES 2016. For both models, only observations without missing values across the response and explanatory variables in each type of model are included. Therefore, the actual number of observations used in these models will be less than 2818.

For the models considered below, the MCMC algorithm will be run for 20000 iterations. The first 10000 iterations are burn-in iterations, whereas the remaining 10000 iterations are used for statistical inference. Similar to the method proposed by Greene (2012) for binary probit and ordinal probit models, two additional calculations are made: (1) the marginal effects of a change in each of the explanatory variables on the probability of an outcome of the response variable, and (2) the predicted probability of an outcome of the response variable for different values of the explanatory variables. The samples from the Gibbs sampling algorithm for the regression coefficients $\boldsymbol{\beta}_\tau$ (and additionally, for ordinal quantile models, the samples for the cut-points $\boldsymbol{\delta}_\tau$) are used for calculating the marginal effects of the explanatory variables. Marginal effects are not calculated for the intercept term.

4.1 Binary quantile models: application

In each of the binary quantile models considered, the following variables are used. The response variable will be whether a voter voted for the Australian Greens in the House of Representatives at the 2016 Federal Election (*green_vote*). Only 2324 responses are included, as these responses contain non-missing values across all the variables listed in the table below. 237 respondents voted for the Greens. 2087 respondents did not vote for the Greens. The variables used are listed in Table 1.

Various binary quantile regression models using the variables in Table 1 are estimated. The 10th quantiles to the 90th quantile at every 10th quantile are chosen. This means the models for $\tau = 0.1$ to $\tau = 0.9$ at increments of 0.1 are chosen. For comparison, the binary probit model will also be considered.

The posterior means for the parameters of the binary probit model and binary quantile model for $\tau = 0.5$ are shown in Table 2. The marginal effects are shown in Table 3. For both models, the regression coefficients are significant as the 95% credible intervals do not include 0. The coefficients show that if the respondent is non-religious or feels that the environment or global warming are extremely important policy issues on deciding how to vote, it is more likely that this person will vote for the Greens. The regression coefficient estimates are different for the probit and quantile models, but they are difficult to interpret. The marginal effects for the quantile model for $\tau = 0.5$ have the following interpretation. A person who believes the environment is extremely important in deciding how to vote is 3.44% ($0.0344\times$ 100%) more likely to vote for the Greens. For a one-year increase in age, the person is 0.14% less likely ($-0.0014 \times 100\%$) to vote for the Greens. The marginal effects for the same variables in the binary probit model are interpreted in a similar way.

As an illustration, the plots for the posterior means and 95% C.I of the marginal effects for the variables *age*, *bachelor*, *bachelor*, *high_tax*, *imp_env*, *imp_warm* and *soc_serv* for the binary quantile models for $\tau = 0.1$ (on the left: QR10) to $\tau = 0.9$ (on the right QR90) are shown in Figure 1. For the plots for binary models, unless otherwise stated, the 10th quantile is denoted as QR10 ('QR' means quantile regression, and 10 means the 10th quantile), and other quantiles are denoted similarly.

The marginal effects of *age* are negative, whereas the marginal effects for other variables are positive. This means that a person, for example, who has a university degree, or feels that the environment is an extremely important, or strongly supports measures to reduce inequality is more likely to vote for the Greens. The marginal effect of each variable for $\tau = 0.9$ is the largest in magnitude.

Variable	Description
green_vote	$Y = 1$ if voted Greens in the House of Representatives; $Y = 0$ otherwise.
age	Age of the respondent in the year 2016. Calculated as 2016 minus the year of birth.
no_religion	Non-religiosity. 1 for non-religious; 0 otherwise.
bachelor	At least a Bachelor's degree. Respondent with at least a bachelor's degree = 1; 0 otherwise.
imp_env	Environment being an extremely important policy issue in deciding how to vote during the election. 1 = extremely important; 0 otherwise.
imp_warm	Global warming being an extremely important policy issue in deciding how to vote during the election. 1 = extremely important; 0 otherwise.
reduce_inequal	Support for reducing income inequality. Score of 1 to 10. 10 strongest. 1 weakest.
high_tax	Support for the statement that high taxes create disincentives to work harder. Score of 1 to 5. 5 strongest. 1 weakest.
soc_serv	Support for an increase in government spending on social services, if the government has to increase social services expenditure or reduce taxes. Score of 1 to 5. 5 strongest. 1 weakest.

Table 1: Description of the variables. Response variable: *green_vote*. Explanatory variables: all other variables.

Model	Variable	Mean	S.D.	95% C.I.
Probit	(Intercept)	−2.472	0.266	[−3.003, −1.958]
	age	−0.019	0.003	[−0.024, −0.015]
	no_religion	0.481	0.086	[0.311,0.648]
	bachelor	0.190	0.087	[0.020,0.362]
	imp_env	0.451	0.119	[0.219,0.684]
	imp_warm	0.434	0.113	[0.216,0.655]
	reduce_inequal	0.074	0.025	[0.025,0.123]
	high_tax	0.128	0.040	[0.052,0.206]
	soc_serv	0.168	0.035	[0.098,0.238]
50th Quantile	(Intercept)	−7.409	0.828	[−9.100, −5.859]
($\tau = 0.5$)	*age*	−0.060	0.007	[−0.073, −0.046]
	no_religion	1.348	0.253	[0.852,1.837]
	bachelor	0.620	0.246	[0.144,1.106]
	imp_env	1.366	0.389	[0.631,2.120]
	imp_warm	1.260	0.358	[0.572,2.005]
	reduce_inequal	0.214	0.077	[0.069,0.370]
	high_tax	0.431	0.122	[0.195,0.670]
	soc_serv	0.594	0.112	[0.381,0.815]

Table 2: Posterior means, standard deviations and 95% C.I. for the parameters in binary models

Model	Variable	Mean	S.D.	95% C.I.
Probit	*age*	−0.0020	0.0003	[−0.0025, −0.0015]
	no_religion	0.0572	0.0121	[0.0344,0.0821]
	bachelor	0.0203	0.0097	[0.0021,0.0402]
	imp_env	0.0489	0.0137	[0.0229,0.0766]
	imp_warm	0.0500	0.0148	[0.0228,0.0804]
	reduce_inequal	0.0074	0.0025	[0.0025,0.0123]
	high_tax	0.0129	0.0041	[0.0052,0.0211]
	soc_serv	0.0169	0.0037	[0.0099,0.0243]
50th Quantile	*age*	−0.0014	0.0002	[−0.0018, −0.0010]
($\tau = 0.5$)	*no_religion*	0.0366	0.0081	[0.0227,0.0553]
	bachelor	0.0154	0.0066	[0.0032,0.0288]
	imp_env	0.0344	0.0105	[0.0148,0.0550]
	imp_warm	0.0333	0.0104	[0.0144,0.0554]
	reduce_inequal	0.0050	0.0018	[0.0016,0.0087]
	high_tax	0.0100	0.0029	[0.0046,0.0157]
	soc_serv	0.0138	0.0027	[0.0089,0.0194]

Table 3: Marginal effects for the explanatory variables in binary models

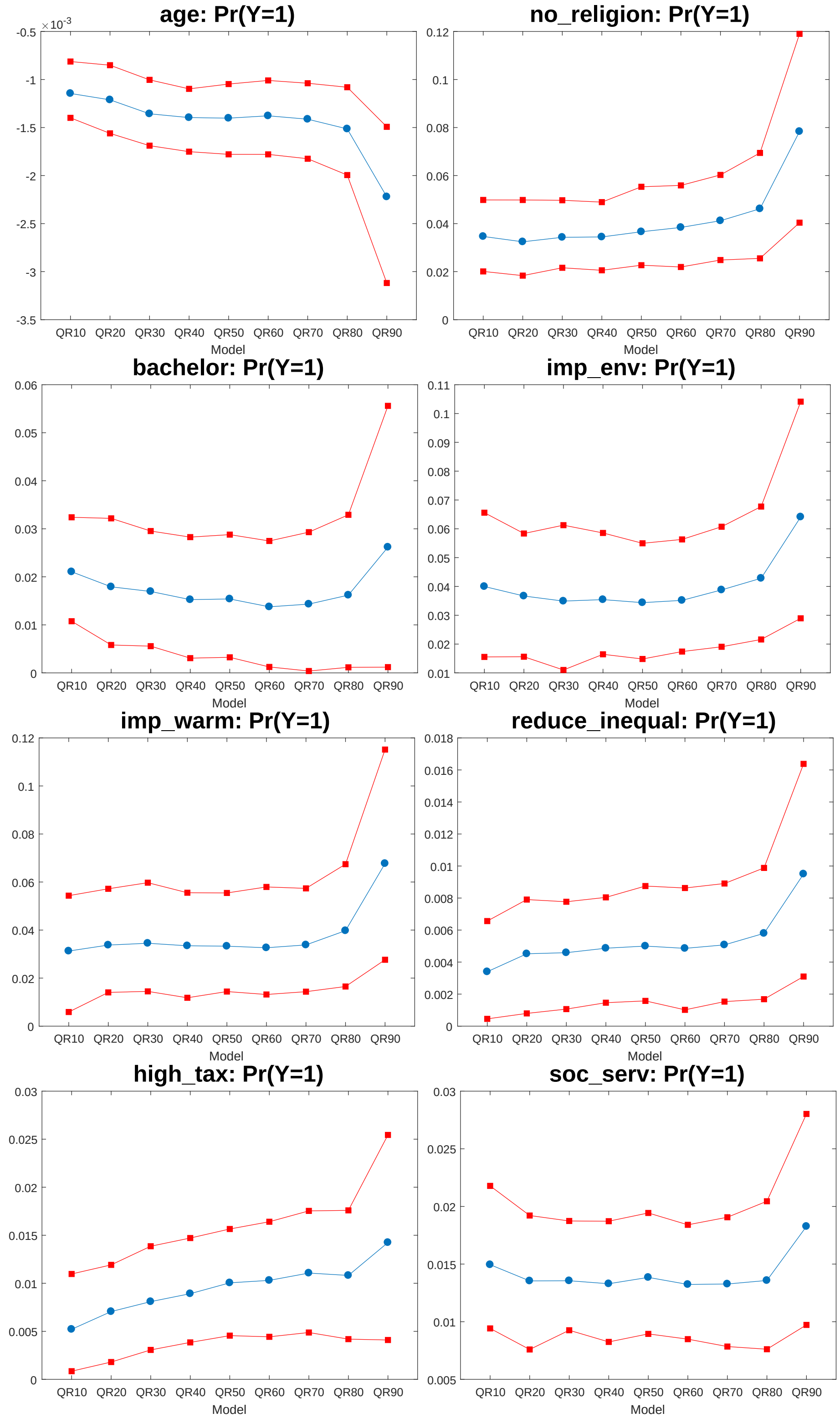

Figure 1: Marginal effects for the binary quantile model coefficients for different quantiles (Posterior mean in blue. 95% upper and lower C.I. in red.)

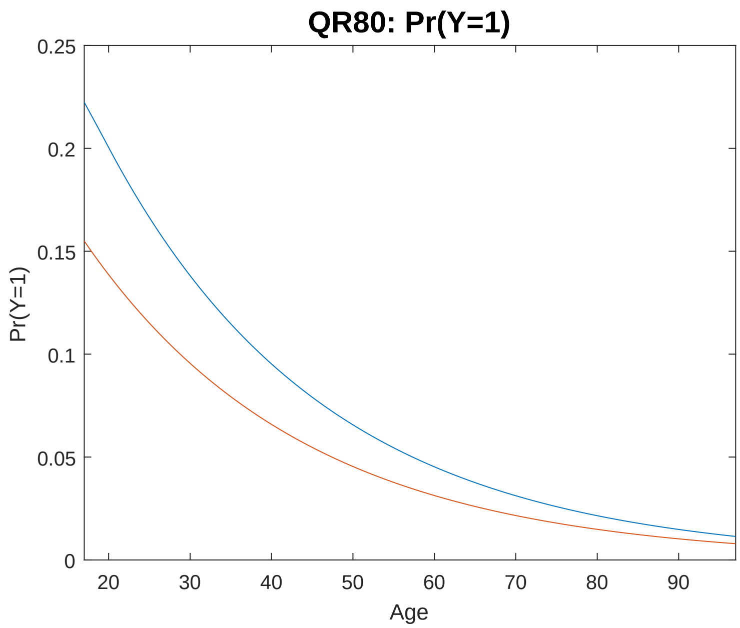

Figure 2: Effect of having at least a bachelor's degree on the predicted probability of a respondent voting for the Greens ($Y = 1$). 80th Quantile ($\tau = 0.8$). (Blue for having at least a bachelor's degree. Red otherwise.)

Figure 2 shows the effect of having at least a bachelor's degree on the predicted probabilities of a respondent voting for the Greens ($Y = 1$) for the quantile model at the 80th quantile ($\tau = 0.8$). If someone has at least a bachelor's degree, the predicted probability is higher than someone without one. Also, as the age increases, the predicted probability decreases, meaning that the older someone is, the less likely it is for the person to vote for the Greens.

4.2 Ordinal quantile models: application

For the ordinal quantile models considered, the response variable will be the level of voters' interest in politics generally (*pol_int*). There are 2173 responses with non-missing values across all variables used in the model. 829 respondents say they have a lot of interest in politics. 927 respondents say they have some interest in politics. 417 respondents say they have not much or no interest in politics. Since the research question is to look at the factors contributing to the level of voters' interest in politics generally, the use of an ordinal quantile model is proposed. The ordinal quantile model is estimated at the 10th quantile to 90th quantile ($\tau = 0.1$ to $\tau = 0.9$), at increments of every 10th quantile. The ordinal probit model is also estimated for comparison. The variables for the ordinal models are listed in Table 4. Since there are three outcomes for the response variable (meaning that $J = 3$), there is only one cut-point $\gamma_{\tau,2}$ to be estimated. This is because for the set of cut-point parameters $\gamma_{\tau,0}$, $\gamma_{\tau,1}$, $\gamma_{\tau,2}$ and $\gamma_{\tau,3}$, the following definitions are used: $\gamma_{\tau,0} = -\infty$, $\gamma_{\tau,1} = 0$ and $\gamma_{\tau,3} = \infty$. The only parameter left to estimate is $\gamma_{\tau,2}$. In Table 5, γ_2 and $\gamma_{0.5,2}$ refer to the cut-points under the ordinal probit model and quantile model for $\tau = 0.5$ respectively.

Variable	Description
pol_int	$Y = 3$ for a lot of interest in politics. $Y = 2$ for some interest in politics. $Y = 1$ for not much or no interest in politics.
pol_attention	Score for respondents' attention to politics on newspapers, television, radio, the internet. Score 1-20. The higher the score, the more attention.
pol_int_camp	Interest in the 2016 Federal election campaign. Score 1-4. The higher the score, the stronger the interest.
difference	Making a difference by voting. Score 1-5. The higher the score, the stronger the belief in making a difference.
sex	Gender. Male = 1; 0 otherwise.
age	Age of the respondent in the year 2016. Calculated as 2016 minus the year of birth.
bachelor	At least a Bachelor's degree. Respondent with at least a bachelor's degree = 1; 0 otherwise.

Table 4: Description of the variables for the ordinal models. Response variable: *pol_int*. Explanatory variables: all other variables.

Model	Variable	Mean	S.D.	95% C.I
Probit	(Intercept)	−4.990	0.181	[−5.344,−4.646]
	pol_attention	0.165	0.014	[0.138,0.193]
	pol_int_camp	1.018	0.048	[0.926,1.113]
	difference	0.129	0.024	[0.083,0.176]
	sex	0.234	0.056	[0.123,0.345]
	age	0.020	0.002	[0.016,0.023]
	bachelor	0.294	0.060	[0.176,0.413]
	γ_2	2.097	0.059	[1.988,2.216]
50th Quantile ($\tau = 0.5$)	(Intercept)	−8.871	0.393	[−9.641, −8.118]
	pol_attention	0.269	0.028	[0.214,0.324]
	pol_int_camp	1.827	0.094	[1.641,2.012]
	difference	0.201	0.049	[0.106,0.298]
	sex	0.438	0.115	[0.211,0.660]
	age	0.032	0.004	[0.025,0.039]
	bachelor	0.474	0.124	[0.224,0.710]
	$\gamma_{0.5,2}$	2.993	0.007	[2.974,2.9998]

Table 5: Posterior means, standard deviations and 95% C.I. for the parameters in ordinal models

Model	Variable	Mean	S.D.	95% C.I.
Probit	*pol_attention*	0.0559	0.0048	[0.0466,0.0655]
	pol_int_camp	0.3448	0.0166	[0.3126,0.3778]
	difference	0.0438	0.0080	[0.0281,0.0596]
	sex	0.0790	0.0191	[0.0417,0.1165]
	age	0.0066	0.0006	[0.0055,0.0079]
	bachelor	0.1014	0.0211	[0.0603,0.1431]
50th Quantile	*pol_attention*	0.0463	0.0052	[0.0363,0.0566]
($\tau = 0.5$)	*pol_int_camp*	0.3140	0.0187	[0.2778,0.3519]
	difference	0.0345	0.0085	[0.0182,0.0514]
	sex	0.0756	0.0202	[0.0365,0.1156]
	age	0.0054	0.0007	[0.0042,0.0068]
	bachelor	0.0842	0.0230	[0.0388,0.1287]

Table 6: Marginal effects for the explanatory variables in ordinal models for the outcome ($Y = 3$)

Table 5 shows the posterior means, standard deviations and 95% C.I. for the regression coefficients and the cut-points for the ordinal probit and quantile model for $\tau = 0.5$. All the variables are statistically significant as the 95% credible intervals for each of the regression coefficients do not include 0. The coefficients for the variables *pol_attention*, *pol_int_camp*, *difference*, *sex*, *age* and *bachelor* are positive for both models. However, the coefficients and cut-points alone are difficult to interpret as they relate to the latent variable rather than the probability of a particular outcome.

The marginal effects for each variable variable for each outcome $Y = 3$, $Y = 2$ and $Y = 1$ are calculated. Table 6 shows the marginal effects of each of the variables for the outcome $Y = 3$. Holding all else constant, a respondent who believed that he/she will make a difference by voting or pays more attention to politics on newspapers, television, radio and on the internet will be more likely to have a lot of interest in politics.

The marginal effects for the quantile model are lower for each of the variables. For example, the marginal effect for *pol_attention* in the ordinal probit model is 0.0559, which means a 1 unit increase in the score for the respondents' attention to politics in the media would increase the probability of having a lot of interest in politics by 5.59%. However, for the quantile model, the marginal effect is 0.0463, which is smaller.

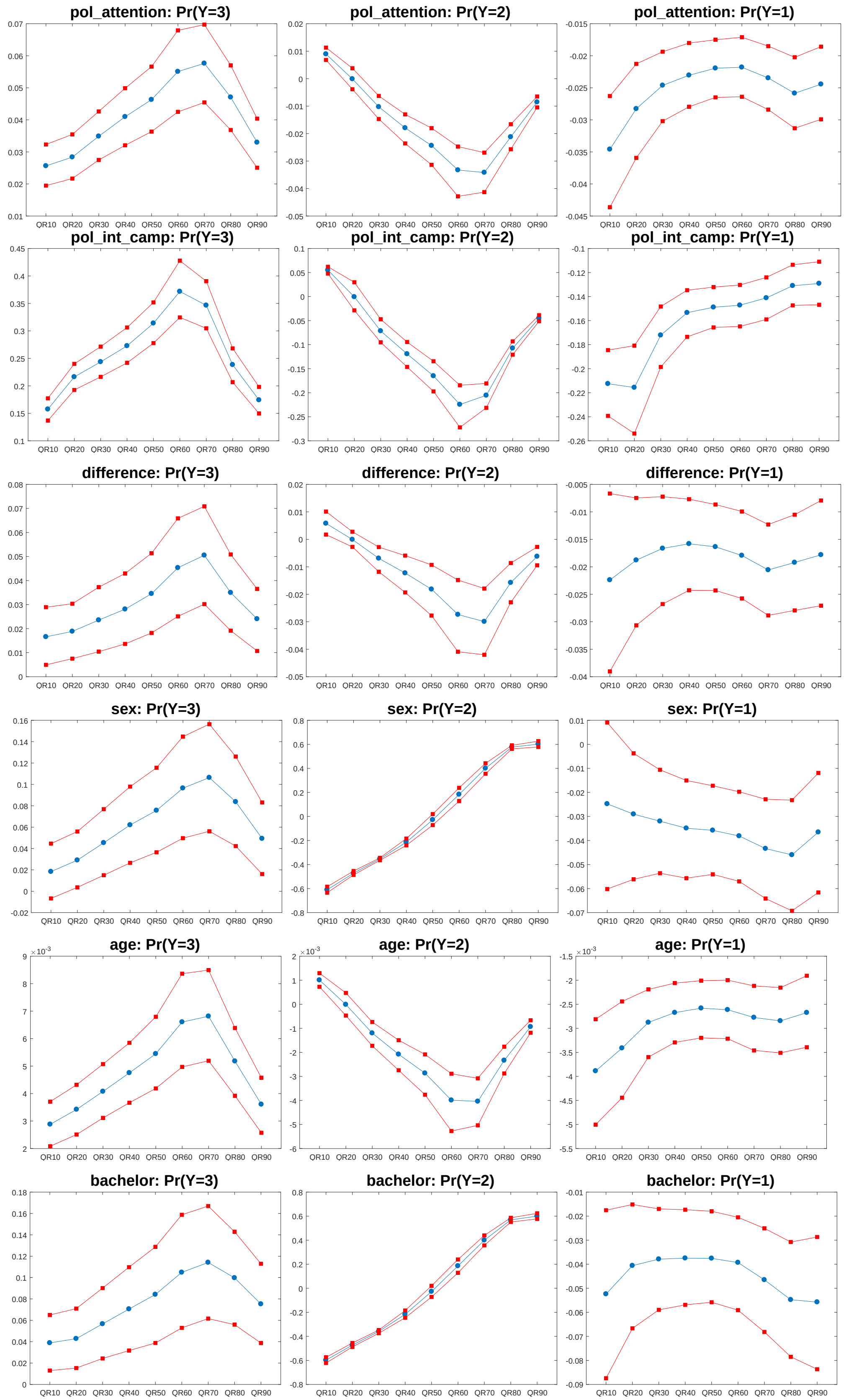

Figure 3: Marginal effects for the ordinal model coefficients for different quantiles (Posterior mean in blue. 95% upper and lower C.I. in red.)

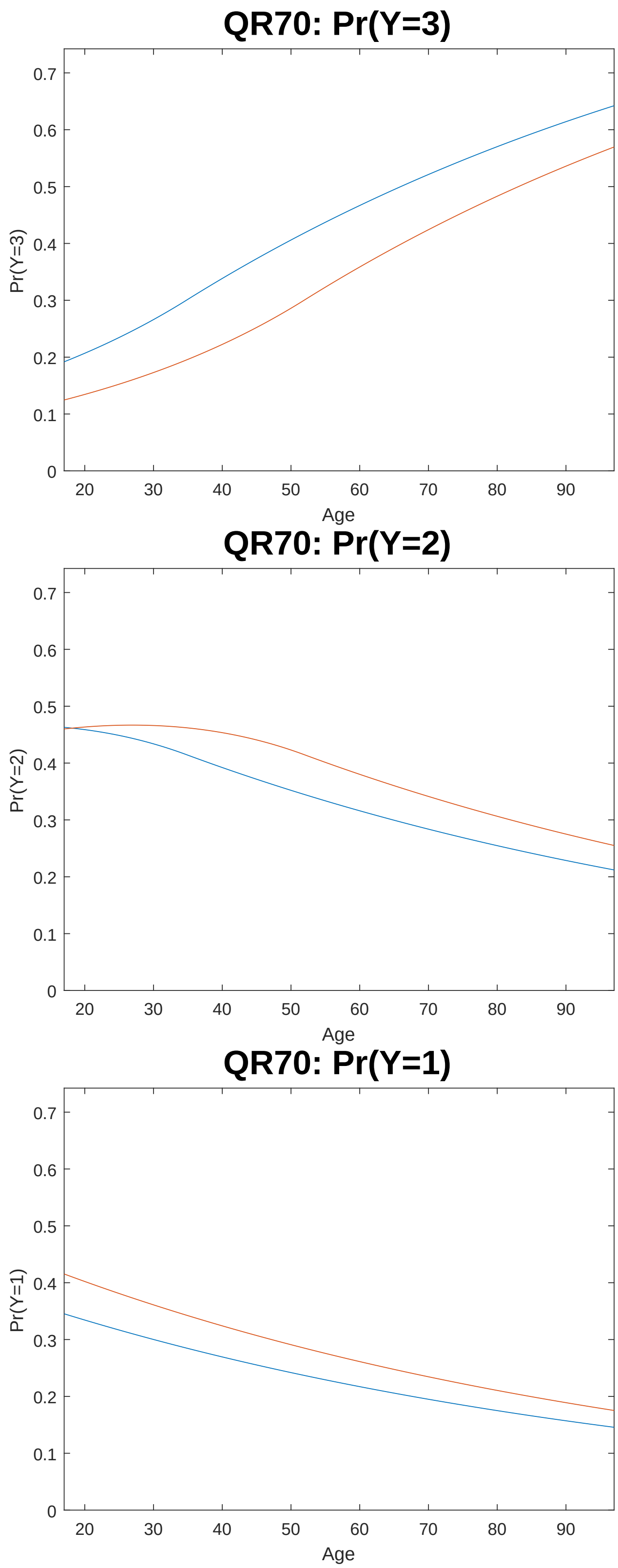

Figure 4: Effect of having at least a bachelor's degree on the predicted probability of having a lot of interest in politics ($Y = 3$), having some interest in politics ($Y = 2$), or having not much or no interest in politics ($Y = 1$). 70th Quantile ($\tau = 0.7$). (Blue for having at least a bachelor's degree. Red otherwise.)

Figure 3 shows the posterior mean of the marginal effect (blue) and the 95% credible intervals (red) for each of the variables for each of the three outcomes $Y = 3$, $Y = 2$ and $Y = 1$. The left-hand side of each plot shows the marginal effects for the 10th quantile (QR10), whereas the right-hand side shows that for the 90th quantile (QR90). Once again, 'QR' means quantile regression and the number denotes the particular quantile. Unless stated otherwise, this will be the notation for the plots for ordinal models.

For the outcome $Y = 3$, the 95% C.I. of most of the variables do not include 0, except for the variable *sex* for the quantile model for $\tau = 0.1$. The marginal effects of *pol_attention*, *pol_int_camp*, *difference*, *age* and *bachelor* are positive. The interpretations for the marginal effects of the variables *pol_attention* and *pol_int_camp* are as follows. Holding all other variables constant, for example, someone who paid more attention to politics on the media or was more interested in the 2016 Federal election campaign is more likely to have a lot of interest in politics generally. Some of the largest magnitudes for the marginal effects for all explanatory variables are quantile models for $\tau = 0.6$ and $\tau = 0.7$, whereas some of the smallest magnitudes are found for $\tau = 0.1$ and $\tau = 0.9$.

For the outcome $Y = 2$, the signs of marginal effects of all variables are different for quantile models of different τ. For the variables *pol_attention*, *pol_int_camp*, *difference* and *age*, the marginal effects are negative for quantile models from at least $\tau = 0.3$. For $\tau = 0.2$, the marginal effects for these variables are statistically insignificant. For $\tau = 0.1$, the marginal effects are positive. However, for the variables *sex* and *bachelor*, the marginal effects are insignificant for $\tau = 0.5$, positive for $\tau > 0.5$ but negative for $\tau < 0.5$.

For the outcome $Y = 1$, the 95% C.I. of marginal effects of most variables cover only negative values across all quantiles, except for the variable *sex* for $\tau = 0.1$. For example, if an individual believed that voting makes a difference, the probability of the person having little to no interest in politics decreases, as the marginal effect for *difference* for this outcome is negative. Similar interpretations could be made for the variable *pol_attention* and *pol_int_camp*.

The effect of having at least a bachelor's degree on the predicted probability of having different levels of interest in politics can be calculated for the ordered quantile model at the 70th quantile ($\tau = 0.7$). This is shown in Figure 4. The predicted probability is calculated for different ages for two different scenarios: a person with at least a bachelor's degree (bachelor = 1, blue curve) and otherwise (bachelor = 0, red curve). An increase in age (holding all else constant) will increase the predicted probability of a lot of interest in politics ($Y = 3$). The predicted probability is higher for someone with at least a bachelor's degree than someone who does not. However, an increase in age will decrease the predicted probability of the other two outcomes: some interest in politics ($Y = 2$) and not much or no interest in politics ($Y = 1$). For the outcome $Y = 1$, the probability is higher for those respondents without at least a bachelor's degree than those with one. This model seems to indicate that, holding all other variables constant, older respondents are more likely to be very interested in politics than younger respondents, and that higher level of education leads to a stronger interest in politics in general.

5. Conclusion

This paper has extended the use of quantile regression models to the Australian Election Studies (AES) survey data. The quantile regression models are used to address two research questions: the impact of factors that lead to a respondent to vote for the Greens in the House of Representatives at the Australian federal election 2016; and the impact of factors that lead to a respondent being more interested in politics in general. This paper proposes using the Bayesian MCMC approach to calculate the marginal effects for the explanatory variables in quantile regressions for binary and ordinal response data. The advantage of this method is that it is easier to interpret the regression coefficients, as marginal effects show the effect of changing each explanatory variable on the probability of each observed outcome. This paper shows an application of this method to the binary and ordinal quantile regressions for the AES and compares the marginal effects for the explanatory variables for different quantiles.

Disclaimer

Researchers who collected the data and carried out the original analysis for the Australian Election Study (AES) 2016 (McAllister et al., 2016) bear no responsibility for the analysis in this presentation and interpretation of the results from this analysis.

Acknowledgments

Charles Au's presentation in JSM 2017 was financially supported by the University of Sydney Postgraduate Research Support Scheme (PRSS) and the Business School Research Travel Support Scheme (RTSS).

References

Albert, J. H. and Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, 88: 669–679.

Andrews, D. F. and Mallows, C. L. (1974). Scale mixtures of normal distributions. *Journal of the Royal Statistical Society . Series B (Methodological)*, 36: 99–102.

Benoit, D. F. and van den Poel, D. (2012). Binary quantile regression: a Bayesian approach based on the asymmetric Laplace distribution. *Journal of Applied Econometrics*, 27: 1174–1188.

Greene, W. H. (2012). Econometric Analysis. 7th ed. (International Edition). England: Pearson.

Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica*, 46: 33–50.

Kozumi, H. and Kobayashi, G. (2011). Gibbs sampling methods for Bayesian quantile regression. *Journal of Statistical Computation and Simulation*, 81: 1565–1578.

Lavín, F. V., Flores, R., and Ibarnegaray, V. (2017). A Bayesian quantile binary regression approach to estimate payments for environmental services. *Environment and Development Economics*, 22: 156–176.

McAllister, I., Makkai, T., Bean, C., and Gibson, R. K. (2016). Australian election study, 2016. [Computer file]. Canberra: Australian Data Archive, The Australian National University.

Michael, J. R., Schucany, W. R., and Haas, R. W. (1976). Generating random variates using transformations with multiple roots. *The American Statistician*, 30: 88–90.

Mollica, C. and Petrella, L. (2016). Bayesian binary quantile regression for the analysis of Bachelor-to-Master transition. *Journal of Applied Statistics*, 44: 2791–2812.

Rahman, M. A. (2016). Bayesian quantile regression for ordinal models. *Bayesian Analysis*, 11: 1–24.

Wichitaksorn, N., Choy, S. T. B., and Gerlach, R. (2014). A generalized class of skew distributions and associated robust quantile regression models. *The Canadian Journal of Statistics*, 42: 579–596.

Yu, K. and Moyeed, R. A. (2001). Bayesian quantile regression. *Statistics & Probability Letters*, 54: 437–447.

Yu, K. and Zhang, J. (2005). A three-parameter asymmetric Laplace distribution and its extension. *Communications in Statistics — Theory and Methods*, 34: 1867–1879.