Binary and Ordinal Probit Regression: Applications to Public Opinion on Marijuana Legalization in the United States

Mohit Batham[a], Soudeh Mirghasemi[b], Manini Ojha[c], Mohammad Arshad Rahman[d,*]

[a] *Walmart, India.*
[b] *Department of Economics, Hofstra University, Hempstead, USA*
[c] *Jindal School of Government and Public Policy, O.P. Jindal Global University, Sonipat, India.*
[d] *Department of Economic Sciences, Indian Institute of Technology Kanpur, India.*

Abstract

This chapter presents an overview of a specific form of limited dependent variable models, namely discrete choice models, where the dependent (response or outcome) variable takes values which are discrete and inherently ordered. Within this setting, the dependent variable may take only two values (such as 0 and 1) giving rise to binary models (e.g., probit and logit) or more than two values (say $j = 1, 2, \ldots, J$, where J is a small integer) giving rise to ordinal models (e.g., ordinal probit and ordinal logit). In these models, the primary goal is to model the probability of responses/outcomes conditional on the covariates. We connect the outcomes of a discrete choice model to the random utility framework in economics, discuss estimation techniques, present the calculation of covariate effects and measures to assess model fitting. Some recent advances in discrete data modeling are also discussed. Following the theoretical overview, we utilize the binary and ordinal models to analyze public opinion on marijuana legalization and the extent of legalization in the United States. All computations are done in MATLAB. We obtain several interesting results including that past use of marijuana, belief about legalization and political partisanship are important factors that shape public opinion.

Keywords: Discrete choice model, logit, marijuana legalization, McFadden's R-squared, Pew Research Center.

*Please address correspondence to Mohammad Arshad Rahman, Department of Economic Sciences, Indian Institute of Technology, Kanpur. Room 409, Engineering Sciences Building, IIT Kanpur, Kanpur 208016, India. Phone: +91 512-259-7010. Fax: +91 512-259-7500. This book chapter was completed and accepted while the corresponding author was working as an Associate Professor in the College of Business, Zayed University, UAE.

Email addresses: mohitbatham15@gmail.com (Mohit Batham), Soudeh.Mirghasemi@hofstra.edu (Soudeh Mirghasemi), mojha@jgu.edu.in (Manini Ojha), marshad@iitk.ac.in, arshadrahman25@gmail.com (Mohammad Arshad Rahman)

1. Introduction

This chapter will discuss settings in which the dependent variable we seek to model takes on a range of values that are restricted, broadly defined as **limited dependent variable models**. Within the class of limited dependent variables, a special case arises when the outcome is no longer a continuous measure but a discrete variable. Such data often arise when individuals make a choice from a set of potential discrete outcomes, thus earning the name **discrete choice** models. The most common case of such models occurs when y is a binary response and takes on the values zero and one, indicating whether or not the event has occurred, giving rise to **binary models**[1]. Consider for example, participation in the labour force, whether or not an individual will buy a vehicle, or whether or not a country is part of free trade agreement. In other cases, y may take on multiple (more than two) discrete values, with no natural ordering. Consider for example, the choice of brand of toothpaste or mode of transportation. These are referred to as **multinomial models**. We refer the readers to Train (2009) for a detailed discussion on multinomial models, their estimation and inference. Further, there could be situations where y takes on multiple (more than two) discrete values that are inherently ordered or ranked. For example, scores attached to opinion on surveys (oppose, neutral, support), classification of educational attainment, or ratings on bonds. These give rise to **ordinal models** or **ordered choice models**. In this chapter, we present binary and ordinal probit models along with a brief sketch on their logit counterparts, namely, binary and ordinal logit models.

Discrete choice models have their foundations in the theory of choice in economics, which itself is inherently related with the random utility model (Luce, 1959; Luce and Suppes, 1965; Marschak, 1960). The random utility framework involves a utility maximizing rational individual whose objective is to choose an alternative from a set of *mutually exclusive* and completely *exhaustive* alternatives. The utilities attached with each alternative are completely known to the decision maker and s/he chooses the same alternative in replications of the experiment. However, to a researcher the utilities are unknown, since s/he only observes a vector of characteristics (such as age, gender, income etc.) of the decision maker, referred to as *representative utility*. This forms the systematic component. The unobserved factors form the stochastic part. The stochastic component

[1]Binary models are special cases of both ordinal and multinomial models with more than two categories.

is assigned a distribution, typically continuous, to make probabilistic statements about the observed choices conditional on the representative utility. The distributional specification implies that there exists a continuous latent random variable (or a continuous latent utility) that underlies the discrete outcomes.

When the set of alternatives or outcomes are inherently ordered or ranked, individual choice of a particular alternative can be associated as the latent variable crossing a particular threshold or cut-point. This latent variable threshold-crossing formulation of the ordered choices elegantly connects individual choice behavior and ordinal data models serve as a useful tool in the estimation process. While the theoretical support relates to choice and random utility theory, the econometric techniques are completely general and applicable when the ordering conditions of the data are met.

To understand the application of discrete choice models, we consider the case of legalization of marijuana in the United States (US). The debate around legalization of marijuana has been an important yet controversial policy issue. Marijuana has been proved to be effective in treatment of several diseases and a wealth of new scientific understanding regarding its medicinal benefits are documented in Berman et al (2004); Wilsey et al (2013); Abrams et al (2003, 2007); Ellis et al (2009); Johnson et al (2010); McAllister et al (2011); Guzmán (2003); Duran et al (2010)[2]. However, despite the medicinal benefits, smoking or consumption of marijuana is not completely benign and may cause harmful effects, especially associated with respiratory illnesses and cognitive development (Kalant, 2004; Polen et al, 1993; Meier et al, 2012).[3] As a result, several surveys have been conducted to assess public opinion on the matter. For the purpose of this chapter, we specifically utilize poll data collected by the Pew Research Center for the periods 2013 and 2014 to demonstrate the application of binary and ordinal probit models. While there is an increasing trend in favor of legalizing marijuana based on public opinion, it is noteworthy to study these specific time periods given that the year 2013 marked the first time in more than four decades that majority of Americans favored legalizing the use of marijuana in the US (Dimock et al, 2013).

[2]The reader is directed to the website www.procon.org for a list of 60 peer-reviewed articles (`http://medicalmarijuana.procon.org/view.resource.php?resourceID=000884`) on the effect of marijuana in treatment of the above mentioned diseases.

[3]A list of peer reviewed articles on the public health consequences of marijuana can be obtained from the Office of National Drug Control Policy. Refer to https://www.whitehouse.gov/ondcp/marijuana.

2. Binary Models

Binary models are designed to deal with situations where the outcome (response or dependent) variable is dichotomous i.e., takes only two values, typically coded as 1 for 'success' and 0 for 'failure'. For example, the application presented in Section 6.1 models the response as a 'success' if an opinion is in favor of legalization and a 'failure' otherwise. While a binary model has a dependent variable that takes discrete values, the model can be conveniently expressed in terms of a continuous latent variable z_i[4] as follows,

$$z_i = x_i'\beta + \epsilon_i, \qquad \epsilon_i \overset{iid}{\sim} N(0,1); \;\; \forall\, i = 1, \cdots, n, \tag{1}$$

where x_i is a $k \times 1$ vector of covariates, β is a $k \times 1$ vector of unknown parameters and n denotes the number of observations. Like most applications, the stochastic term ϵ is assumed to be *independently and identically distributed* (*iid*) as a standard normal distribution, i.e., $\epsilon_i \overset{iid}{\sim} N(0,1)$ for $i = 1, \ldots, n$; which gives rise to a *binary probit model*. The latent variable z_i is related to the observed discrete response y_i as follows:

$$y_i = \begin{cases} 1 & \text{if } z_i > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

With only two responses, there is a single cut-point which is typically fixed at 0 for the sake of simplicity. A pictorial representation of the binary outcome probabilities for marijuana legalization is shown in Figure 1. The figure also shows that the cut-point γ_1 is fixed at 0 to anchor the location of the distribution. Besides, the scale is fixed by assuming the variance of the normal distribution is 1. Both the restrictions are necessary to identify the model parameters (see Jeliazkov and Rahman, 2012, for further details).

The likelihood for binary probit model can be expressed as,

$$L(\beta; y) = \prod_{i=1}^{n} \left\{ \Pr(y_i = 0|x_i'\beta)^{(1-y_i)} \Pr(y_i = 1|x_i'\beta)^{y_i} \right\} = \prod_{i=1}^{n} \left\{ \Phi(-x_i'\beta)^{(1-y_i)} \Phi(x_i'\beta)^{y_i} \right\}, \tag{3}$$

where $\Phi(\cdot)$ denotes the *cumulative distribution function (cdf)* of a standard normal distribution.

[4]The continuous latent variable can be interpreted as the difference between utilities from choice 1 and 0 i.e., $z_i = U_{i1} - U_{i0}$, where U denotes utility (Jeliazkov and Rahman, 2012).

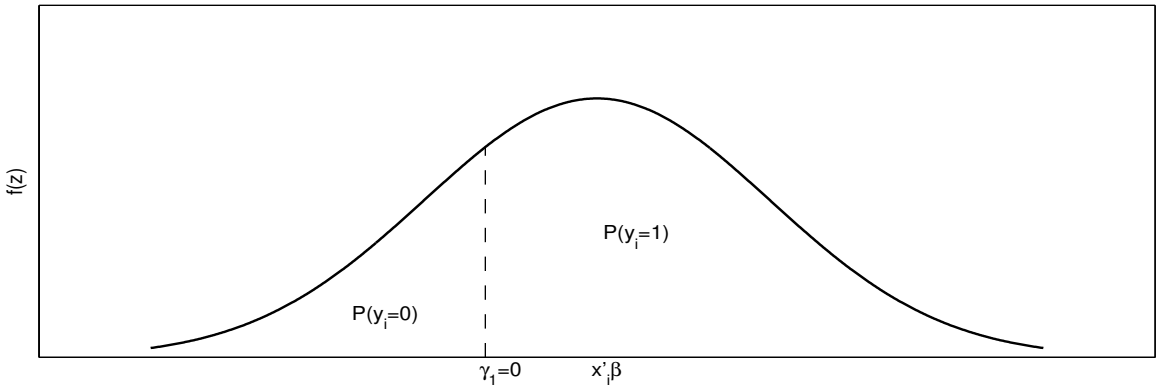

Figure 1: The cut-point γ_1 divides the area under the curve into two parts, the probability of failure and probability of success. In our study, $\mathrm{P}(y_i = 0)$ and $\mathrm{P}(y_i = 1)$ correspond to probability of opposing and supporting marijuana legalization, respectively. Note that for each individual i the mean $x_i'\beta$ and hence the probabilities, $\mathrm{P}(y_i = 0)$ and $\mathrm{P}(y_i = 1)$, will be different. Source: authors' creation.

Given the likelihood in equation 3, the parameters β are estimated by maximizing the logarithm of the likelihood (log-likelihood) using numerical techniques such as the Newton-Raphson method or the BHHH procedure (Train, 2009). The principle behind maximizing the likelihood – known as the maximum likelihood (ML) estimation – is to obtain those parameter values that are most probable to have produced the data under the assumed statistical model. Note that it is convenient to work with the log-likelihood since logarithm being a monotonic function, the maximum of log-likelihood and the likelihood occur at the same parameter values. Once the parameter estimates are available, we may calculate post-estimation constructs such as covariate effects and predicted probabilities. Measures for goodness of fit can also be calculated to assess model fitting. Readers interested in further details about binary data modeling may look into Johnson and Albert (2000, Chap. 3).

Thus far, we have described the binary probit model but the framework can be transformed into a binary logit model (or simply logit model) with the modification that the errors follow a logistic distribution (Hosmer et al, 2013). Like the normal distribution, the logistic distribution is symmetric but has heavier tails relative to a normal distribution. Both the location and scale restrictions apply to the logit model as before, but note that the variance is now fixed at $\pi^2/3$ as compared to 1 in a probit model. To obtain the logit likelihood, the normal *cdf* $\Phi(x_i'\beta)$ in equation (3) is replaced by the logistic *cdf*: $\Lambda(x_i'\beta) = \exp(x_i'\beta)/\big[1 + \exp(x_i'\beta)\big]$. We maximize the resulting log-likelihood to obtain the parameter estimates for the logit model.

The logit model is appealing to researchers in many fields (including epidemiologists) because

of the ease in interpreting its slope coefficient. To see this, let $\theta_i = \Pr(y_i = 1|x_i'\beta)$ denote the probability of success and $x_{.,l}$ be a continuous covariate (or independent variable). Then the logarithm of the odds (log-odds) of success can be expressed as,

$$\log\left(\frac{\theta_i}{1-\theta_i}\right) = x_i'\beta = x_{i,l}\,\beta_l + x_{i,-l}'\,\beta_{-l}.$$

where $x_i = (x_{i,l},\, x_{i,-l})$, $\beta = (\beta_l,\, \beta_{-l})$, and $-l$ in the subscript denotes all covariates/parameters except the l-th covariate/parameter. If we differentiate the log-odds with respect to the l-th covariate, we obtain β_l. Therefore, the slope coefficient β_l represents the log-odds for a 1 unit change in the l-th covariate.

Similarly, the coefficient of an indicator variable (dummy or dichotomous variable) has an interesting interpretation. Let $x_{.,m}$ be an indicator variable, θ_i^1 be the probability of success when $x_{i,m} = 1$, θ_i^0 be the probability of success when $x_{i,m} = 0$. Our goal is to find the expression for the odds-ratio, which measures the odds of success among those with $x_{i,m} = 1$ compared to those with $x_{i,m} = 0$. Then the logarithm of the odds-ratio is,

$$\log\left(\frac{\theta_i^1/(1-\theta_i^1)}{\theta_i^0/(1-\theta_i^0)}\right) = \beta_m + x_{i,-m}'\,\beta_{-m} - x_{i,-m}'\,\beta_{-m} = \beta_m.$$

The odds-ratio is better understood with the help of an example. Suppose y denotes the presence or absence of a heart disease and $x_{.,m}$ denotes whether the person is a smoker or non-smoker. Then, an odds-ratio = 2 implies that heart disease is twice as likely to occur among smokers as compared to non-smokers for the population under study.

3. Ordinal Models

Ordinal regression models generalize the binary models by allowing the dependent variable to have more than two outcomes which are inherently ordered or ranked. Each outcome or category is assigned a score (value or number) with the characteristic that the scores have an ordinal meaning but hold no cardinal interpretation. Therefore, the difference between categories is not directly comparable. For example, the study presented in Section 6.2 codifies the public response to marijuana legalization as follows: 1 for 'oppose legalization', 2 for 'legal only for medicinal use', and 3

for 'legal for personal use'. Here, a score of 2 implies more support for legalization as compared to 1, but we cannot interpret a score of 2 as twice the support compared to a score of 1. Similarly, the difference in support between 2 and 1 is not the same as that between 3 and 2.

Similar to binary models, the ordinal regression model can be conveniently expressed in terms of a continuous latent variable z_i[5] as follows:

$$z_i = x_i'\beta + \epsilon_i, \qquad \forall\, i = 1, \cdots, n, \tag{4}$$

where the notations are identical to that of a binary model. If we assume that the errors are *iid* as a standard normal distribution, i.e., $\epsilon_i \overset{iid}{\sim} N(0,1)$ for $i = 1, \ldots, n$; then we have an *ordinal probit model* (also known as *ordered probit model*). The latent variable z_i is related to the observed discrete response y_i as follows:

$$\gamma_{j-1} < z_i \leq \gamma_j \;\Rightarrow\; y_i = j, \qquad \forall\, i = 1, \cdots, n;\; j = 1, \cdots, J, \tag{5}$$

where $-\infty = \gamma_0 < \gamma_1 \cdots < \gamma_{J-1} < \gamma_J = \infty$ are the cut-points (or thresholds) and y_i is assumed to have J categories or outcomes. A visual representation of the outcome probabilities (for the case of marijuana legalization) and the cut-points are presented in Figure 2. One may observe from Figure 2 that different combinations of (β, γ) can produce the same outcome probabilities giving rise to parameter identification problem. We therefore need to anchor the location and scale of the distribution to identify the model parameters. The former is achieved by setting $\gamma_1 = 0$ and the latter by assuming $\text{var}(\epsilon_i) = 1$. Other identification schemes are possible and the reader is referred to Jeliazkov et al (2008) and Jeliazkov and Rahman (2012) for details.

Given a data vector $y = (y_1, \cdots, y_n)'$, the likelihood for the ordinal probit model expressed as a function of unknown parameters (β, γ) is the following,

$$L(\beta, \gamma; y) = \prod_{i=1}^{n}\prod_{j=1}^{J} \Pr(y_i = j|\beta, \gamma)^{I(y_i=j)} = \prod_{i=1}^{n}\prod_{j=1}^{J} \Big[\Phi(\gamma_j - x_i'\beta) - \Phi(\gamma_{j-1} - x_i'\beta)\Big]^{I(y_i=j)}, \tag{6}$$

[5]The continuous latent construct may represent underlying latent utility, some kind of propensity, or strength of preference (Greene and Hensher, 2010).

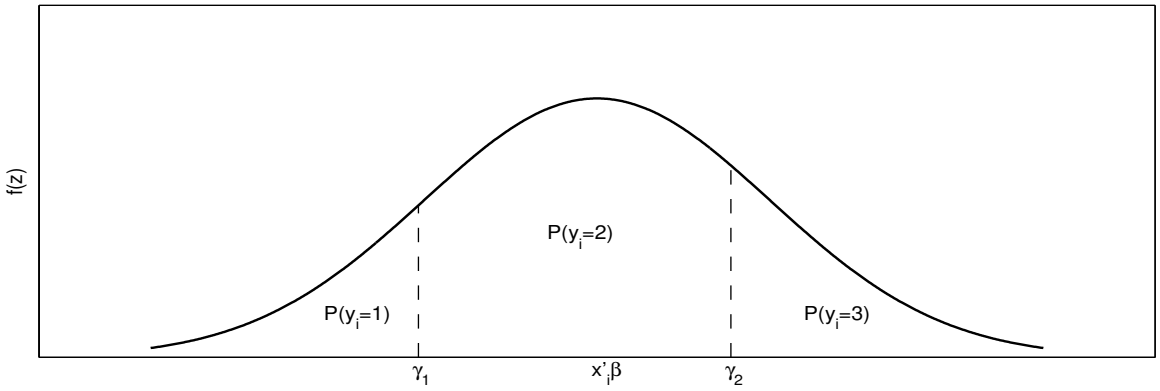

Figure 2: The two cut-points (γ_1, γ_2) divide the area under the curve into three parts, with each part representing the probability of a response falling in the three response categories. The three probabilities $\mathrm{P}(y_i = 1)$, $\mathrm{P}(y_i = 2)$ and $\mathrm{P}(y_i = 3)$ correspond to 'oppose legalization', 'legal only for medical use' and 'legal for personal use', respectively. Note that for each individual i the mean $x_i'\beta$ will be different and so will be the category probabilities. Source: authors' creation.

where $\Phi(\cdot)$ is the *cdf* of a standard normal distribution and $I(y_i = j)$ is an indicator function, which equals 1 if the condition within parenthesis is true and 0 otherwise. The parameter estimates for (β, γ) are obtained by maximizing the log-likelihood, i.e., logarithm of the likelihood given by equation equation (6), either using the Newton-Raphson method or BHHH procedure (Train, 2009). Once the parameter estimates are available, they may be used to calculate the covariate effects, make predictions or assess model fitting. Interested readers may look into Greene and Hensher (2010) or (Johnson and Albert, 2000, Chap. 4) for a detailed review of ordinal data modeling.

Similar to binary models, the framework for ordinal probit model can be transformed into an *ordinal logit model* (or *ordered logit model*) by simply assuming that the error follows a logistic distribution (McKelvey and Zavoina, 1975; McCullagh, 1980). Therefore, for the model in equation (4), we now assume that $\epsilon_i \sim L(0, 1)$ for $i = 1, \ldots, n$, where L denotes a logistic distribution with mean 0 and variance $\pi^2/3$. The likelihood for the ordinal logit model has the same structure as equation (6) with $\Phi(w)$ replaced by $\Lambda(w) = \exp(w)/[1+\exp(w)]$, where w is the argument inside the parenthesis. Analogous to the ordinal probit model, the parameters are estimated using the ML technique.

An interesting property of ordinal logit model is that the ratio of odds of not exceeding a certain category (say j) for any two individuals is constant across response categories. This earns it the name *proportional odds model*. To see this property in effect, let $\theta_{ij} = \Pr(y_i \leq j)$ denote the

cumulative probability that individual i chooses category j or below. For the ordinal logit model, this implies: $\theta_{ij} = \exp(\gamma_j - x_i'\beta)/\big[1 + \exp(\gamma_j - x_i'\beta)\big]$, and $\theta_{ij}/(1-\theta_{ij}) = \exp\big[\gamma_j - x_i'\beta\big]$, where the latter represents the odds for the event $y_i \leq j$. Accordingly, for any two individuals (say 1 and 2), the ratio of odds is,

$$\frac{\theta_{1j}/(1-\theta_{1j})}{\theta_{2j}/(1-\theta_{2j})} = \exp\big[-(x_1 - x_2)'\beta\big]. \tag{7}$$

The odds ratio presented in equation (7) does not depend on the response category j and is proportional to $(x_1 - x_2)$ with β being the constant of proportionality.

4. Covariate Effects and Model Fitting

In binary and ordinal models, the coefficients do not give covariate effects because the link function is non-linear and non-monotonic. Consequently, we need to calculate the covariate effect for each outcome. Let $x_{.,l}$ be a continuous covariate, then the covariate effect for the i-th observation (or individual) in an ordinal probit model is calculated as,

$$\begin{aligned}\frac{\partial \Pr(y_i = j)}{\partial x_{i,l}} &= -\beta_l \left[\phi(\gamma_j - x_i'\beta) - \phi(\gamma_{j-1} - x_i'\beta)\right] \\ &\simeq -\hat{\beta}_l \left[\phi(\hat{\gamma}_j - x_i'\hat{\beta}) - \phi(\hat{\gamma}_{j-1} - x_i'\hat{\beta})\right],\end{aligned} \tag{8}$$

where $\phi(\cdot)$ denotes the probability density function (*pdf*) of a standard normal distribution and $(\hat{\beta}, \hat{\gamma})$ are the ML estimates of the parameters (β, γ). The average covariate effect is computed by averaging the covariate effect in equation (8) across all observations. If the covariate is an indicator variable (say $x_{.,m}$), then the covariate effect for the i-th observation on outcome j $(= 1, \ldots, J)$ is calculated as,

$$\begin{aligned}&\Pr(y_i = j | x_{i,-m}, x_{i,m} = 1) - \Pr(y_i = j | x_{i,-m}, x_{i,m} = 0) \\ &= \big[\Phi(\gamma_j - x'^{\dagger}_i\beta) - \Phi(\gamma_{j-1} - x'^{\dagger}_i\beta)\big] - \big[\Phi(\gamma_j - x'^{\ddagger}_i\beta) - \Phi(\gamma_{j-1} - x'^{\ddagger}_i\beta)\big] \\ &\simeq \big[\Phi(\hat{\gamma}_j - x'^{\dagger}_i\hat{\beta}) - \Phi(\hat{\gamma}_{j-1} - x'^{\dagger}_i\hat{\beta})\big] - \big[\Phi(\hat{\gamma}_j - x'^{\ddagger}_i\hat{\beta}) - \Phi(\hat{\gamma}_{j-1} - x'^{\ddagger}_i\hat{\beta})\big],\end{aligned} \tag{9}$$

where $x_i^{\dagger} = (x_{i,-m},\, x_{i,m} = 1)$ and $x_i^{\ddagger} = (x_{i,-m},\, x_{i,m} = 0)$. The average covariate effect is calculated by averaging the covariate effect given in equation (9) across all observations. Note that for ordinal models, the sign of the regression coefficient translates unambiguously into the sign of covariate

effect only for the lowest and highest categories of the response variable. Covariate effect for the middle categories cannot be known *a priori*.

For the binary probit models, the expressions for covariate effects simplify because with only two outcomes there is a single cut-point which is fixed at 0. For a continuous variable, the covariate effect is given by the expression,

$$\frac{\partial \Pr(y_i = 1)}{\partial x_{i,l}} = \beta_l \, \phi(x_i'\beta) \simeq \hat{\beta}_l \, \phi(x_i'\hat{\beta}), \tag{10}$$

and the same for an indicator variable is given by the expression,

$$\begin{aligned} &\Pr(y_i = 1|x_{i,-m}, x_{i,m} = 1) - \Pr(y_i = 1|x_{i,-m}, x_{i,m} = 0) \\ &= \Phi(x'^{\dagger}_i\beta) - \Phi(x'^{\ddagger}_i\beta) \simeq \Phi(x'^{\dagger}_i\hat{\beta}) - \Phi(x'^{\ddagger}_i\beta), \end{aligned} \tag{11}$$

where all the notations have been explained in the previous paragraph. Once again, the average covariate effect is computed by averaging across all observations. While the discussion on covariate effects has considered binary and ordinal probit models because of their implementation in the applications, covariate effects for binary and ordinal logit models can be calculated analogously by replacing the normal *pdf*'s and *cdf*'s with the logistic *pdf*'s and *cdf*'s at appropriate places.

To assess the goodness of model fit, we calculate three measures: likelihood ratio (LR) test statistic, McFadden's R-squared (McFadden, 1974) and hit-rate (Johnson and Albert, 2000). For the null hypothesis $H_0 : \beta_2 = \ldots = \beta_k = 0$, the LR test statistic λ_{LR} is defined as follows:

$$\lambda_{LR} = -2[\ln L_0 - \ln L_{\text{fit}}] \quad \overset{H_0}{\sim} \quad \chi^2_{k-1}, \tag{12}$$

where $\ln L_{\text{fit}}$ is the log-likelihood of the fitted model and $\ln L_0$ is the log-likelihood of the intercept-only model. Under the null hypothesis, λ_{LR} follows a chi-square distribution with degrees of freedom equal to $k-1$, i.e., the number of restrictions under the null hypothesis. So, we calculate the statistic λ_{LR} and compare it with χ^2_{k-1} for a given level of significance. If $\lambda_{LR} > \chi^2_{k-1}$, then we reject the null hypothesis. Otherwise, we do not reject the null hypothesis.

Another popular goodness of fit measure for discrete choice models is McFadden's R-squared (R^2_M), due to McFadden (1974). The McFadden's R-squared, also referred to as pseudo R-squared

or likelihood ratio index, is defined as follows,

$$R^2_M = 1 - \frac{\ln L_{\text{fit}}}{\ln L_0}. \tag{13}$$

The R^2_M is intuitively appealing because it is bounded between 0 and 1, similar to the coefficient of determination (R^2) in linear regression models. When all slope coefficients are zero, the R^2_M equals zero; but in discrete choice models R^2_M can never equal 1, although it can come close to 1. While higher values of R^2_M implies better fit, the value as such has no natural interpretation in sharp contrast to R^2 which denotes the proportion of variation in the dependent variable explained by the covariates.

While both LR test statistic and McFadden's R-squared are commonly used in applied studies, the hit-rate is relatively uncommon. The hit-rate (HR) is defined as the percentage of correct predictions i.e., percentage of observations for which the model correctly assigns the highest probability to the observed response category. Mathematically, the HR can be defined as follows,

$$HR = \frac{1}{n}\sum_{i=1}^{n} I\Bigg(\Big(\max_j \{\hat{p}_{ij}\}_{j=1}^{J}\Big) = y_i\Bigg), \tag{14}$$

where $\hat{p}_{ij}$ is the predicted probability that individual i selects outcome j, and $I(\cdot)$ is the indicator function as defined earlier.

5. Some Advances in Discrete Choice Modeling

Till now, we have looked at binary and ordinal models and their Classical (or Frequentist) approach to estimation via the maximum likelihood technique, which only involves the likelihood function (say $L(\theta; y) \equiv f(y|\theta)$, where θ is the parameter vector). The Classical approach assumes that model parameters are unknown but have fixed values and hence the parameters cannot be treated as random variables. In contrast, Bayesian approach to estimation utilizes the Bayes' theorem,

$$\pi(\theta|y) = \frac{f(y|\theta)\pi(\theta)}{\int f(y|\theta)\pi(\theta)\, d\theta},$$

to update the belief/information about θ (considered a random variable) by combining information from the observed sample (via the likelihood function) and non-sample or prior beliefs (arising

from previous studies, theoretical consideration, researcher's belief, etc.) represented by the prior distribution $\pi(\theta)$. Inference is based on the posterior distribution $\pi(\theta|y)$. Bayesian approach provides several advantages including finite sample inference, working with likelihoods which are difficult to evaluate, and advantages in computation. Interested readers may look into Greenberg (2012) for details on the Bayesian approach and Poirier (1995) for a comparison of Classical and Bayesian estimation methods.

The posterior density for binary and ordinal models do not have a tractable density and so the parameters cannot be sampled directly. While the Metropolis-Hastings (MH) algorithm (Metropolis et al, 1953; Hastings, 1970) can be employed to sample the parameters, the standard and more convenient approach is to consider *data augmentation* (Tanner and Wong, 1987). In this approach, the joint posterior density is augmented by a latent variable z and the augmented joint posterior $\pi(\theta, z|y)$ is written as,

$$\pi(\theta, z|y) \propto \pi(\theta) f(z, y|\theta) = \pi(\theta) f(z|\theta) f(y|z, \theta).$$

For a binary probit model, $\theta = \beta$ and $f(y|z, \beta) \equiv f(y|z)$; whereas for an ordinal probit model $\theta = (\beta, \gamma)$ and $f(y|z, \beta, \gamma) \equiv f(y|z, \gamma)$. The two equivalencies arise because, given a latent observation z_i, y_i is known with certainty regardless of β for $i = 1, \ldots, n$. This de-linking of the likelihood function from β, made possible through data augmentation, simplifies the estimation procedure and allows sampling of β through a Gibbs process (Geman and Geman, 1984) – a well known Markov chain Monte Carlo (MCMC) technique. The latent variable z is sampled element-wise from a truncated normal distribution. For ordinal models, a monotone transformation of the cut-points, γ, is sampled using an MH algorithm. The MCMC algorithms for estimating binary and ordinal probit models outlined here were introduced in Albert and Chib (1993). Other notable references that describe the Bayesian modeling and estimation of binary and ordinal outcomes in great detail include Johnson and Albert (2000), Greenberg (2012), and Jeliazkov and Rahman (2012). Bayesian estimation of logit model is based on the same principle and presented in Holmes and Held (2006) and Jeliazkov and Rahman (2012).

The binary and ordinal models considered in this chapter, whether estimated using the Classical or the Bayesian techniques, provide information on the average probability of outcomes conditional

on the covariates. However, interest on quantiles of the response variable as a robust alternative to mean regression have grown enormously since the introduction of quantile regression in Koenker and Bassett (1978). Quantile modeling gained further momentum with the development of Bayesian quantile regression by Yu and Moyeed (2001), where the authors create a working likelihood by assuming that the errors follow an asymmetric Laplace (AL) distribution (Yu and Zhang, 2005). Binary quantile regression was proposed by Kordas (2006) and its Bayesian formulation was presented by Benoit and Poel (2010). Rahman (2016) introduced Bayesian quantile regression with ordinal responses and estimated the model using MCMC techniques. The corresponding R package *bqror* for estimating the quantile ordinal model is described in Maheshwari and Rahman (2023) along with the computation of marginal likelihood for comparing alternative quantile models. A flexible form of Bayesian ordinal quantile regression was proposed in Rahman and Karnawat (2019). Some recent research on binary quantile regression in the panel/longitudinal set up include Rahman and Vossmeyer (2019), and Bresson et al (2021). Two applied studies employing binary and ordinal quantile framework are Ojha and Rahman (2021) and Omata et al (2017), respectively. Interested readers may explore the above mentioned papers and references therein to develop a thorough understanding on binary and ordinal quantile modeling and their applications.

6. Applications: Public Opinion on Legalization of Marijuana in the United States

In the US, marijuana is illegal under the federal law as per the Controlled Substances Act of 1970. The Act classifies marijuana as a schedule I drug i.e., a drug with no accepted medical value, high potential for abuse and not safe to use even under medical supervision (Drug Enforcement Administration, 2011). However, state laws pertaining to marijuana have evolved overtime. Up until 2016, 29 states had either legalized, allowed access for medical reasons or decriminalized its use (See Figure 3). More legalization efforts are appearing in the remaining states of US. Such revisions in state laws represent a change in public attitude that is aptly reflected in survey data collected by independent polling agencies such as the Pew Research Center, General Social Survey and Gallup. Figure 4 shows an increasing trend in favor of legalizing marijuana based on public opinion. Besides, political standing on marijuana has also popularized the debate on its legalization and may have affected public opinion regarding it. The gradual growth in support of marijuana perhaps signals greater better public insight on the medicinal value of marijuana (Earleywine, 2005) and the social

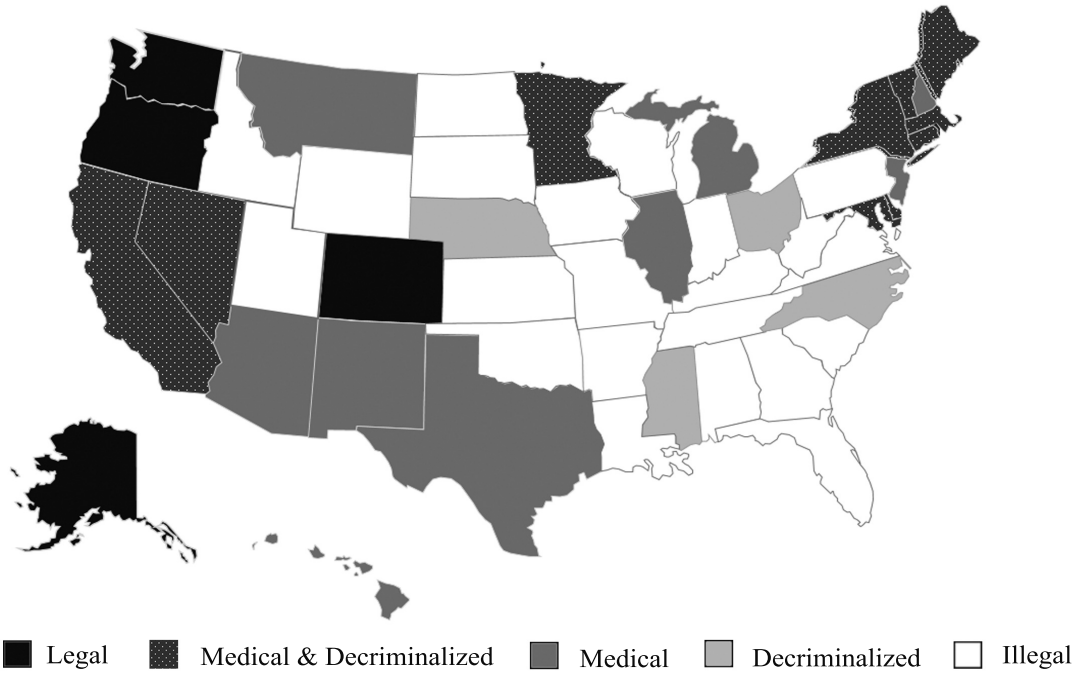

Figure 3: Marijuana state laws in the US as of 6th January, 2016. Source: authors' creation from various sources.

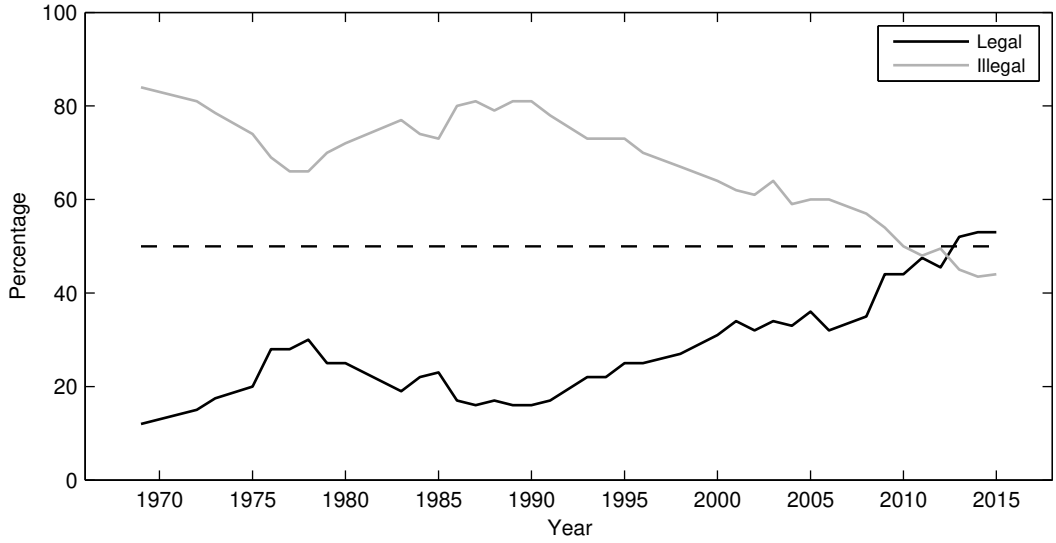

Figure 4: Public opinion on marijuana legalization for the period 1969-2015. The black dashed line is the 50 percent benchmark. Data source: Pew Research Center, General Social Survey and Gallup. We have averaged the percentages for years with multiple surveys. The combined percentage of the two opinions is below 100 since on average 4 percent of respondents answered "don't know" or "refused to answer".

cost of prohibition that includes illegal trade, racially skewed arrests of African Americans and huge enforcement cost (Shepard and Blackley, 2007).

While policies on marijuana use are in the early stages of formulation as states evaluate its costs

and benefits (Winterbourne, 2012), more states decriminalizing its use may cause a major policy shift at the federal level (Ferner, 2015). In this regard, some scholars argue that public policies ought to be guided by public opinion such that mass opinion and democracy is upheld (Monroe, 1998; Paletz et al, 2015). Besides, Shapiro (2011) cites a large number of studies to argue that public opinion influences government policy making in the US. Therefore, it is imperative to study and identify the factors that significantly impact US public opinion towards marijuana legalization.

In the next section, we employ a binary probit model to analyze public opinion on marijuana legalization and thereafter implement the ordinal probit model to analyze public opinion on the extent of marijuana legalization. The choice of probit models over their logit counterparts is driven by practical considerations – probit models are tractable in univariate cases and can be generalized to multivariate and hierarchical settings (Jeliazkov et al, 2008). In contrast, logistic model based on logistic distribution cannot model correlations in multivariate settings.

6.1. Binary Probit Model - Marijuana Legalization

In this subsection, we summarize the March 2013 Political Survey data used for the study, estimate a binary probit model, present the results, model fitting diagnostics, and covariate effects.

6.1.1. Data

We utilize the March 2013 Political Survey data from the Pew Research Center to analyze public opinion on marijuana and identify the factors that significantly impact the probability of supporting its legalization. The survey was conducted during the period March 13-17, 2013, by Abt SRBI (Schulman, Ronca & Buculvas, Inc) for the Pew Research Center for the People and the Press. The survey selected and interviewed a representative sample of 1,501 adults living in the US. Of the 1501 adults, 750 individuals were interviewed over land line and the remaining 751 individuals over cell phone. The available sample had several respondents with missing values ("don't know" or "refused to answer") on the variables of interest, along with 49 respondents who were unsure about marijuana legalization. After removing data on these respondents, we have a sample of 1182 observations available for the study[6].

[6]The data file ("Mar13Data.xlsx") and Matlab codes file ("ProbitScript.m") for estimation and post-estimation constructs can be downloaded from the corresponding author's webpage: `https://www.arshadrahman.com/Research.html`.

Table 1: Descriptive summary of the variables (March 2013 Political Survey).

VARIABLE		MEAN	STD
LOG AGE		3.86	0.40
LOG INCOME		10.64	0.98
HOUSEHOLD SIZE		2.72	1.44
	CATEGORY	COUNTS	PERCENTAGE
PAST USE		554	46.87
MALE		570	48.22
EDUCATION	BACHELORS & ABOVE	426	36.04
	BELOW BACHELORS	360	30.46
	HIGH SCHOOL & BELOW	396	33.50
TOLERANT STATES		374	31.64
RACE	WHITE	938	79.36
	AFRICAN AMERICAN	142	12.01
	OTHER RACES	102	8.63
PARTY AFFILIATION	REPUBLICAN	353	29.86
	DEMOCRAT	404	34.18
	INDEPENDENT & OTHERS	425	35.96
RELIGION	PROTESTANT	494	41.79
	ROMAN CATHOLIC	258	21.83
	CHRISTIAN	138	11.68
	CONSERVATIVE	72	6.09
	LIBERAL	220	18.61
PUBLIC OPINION	FAVOR LEGALIZATION	622	52.62
	OPPOSE LEGALIZATION	560	47.38

In this application, the dependent variable is the response to the question: "Do you think the use of marijuana should be made legal, or not?". The responses were recorded as 'yes, legal' (i.e. favor legalization), 'no, illegal' (i.e., oppose legalization) or 'don't know or refused'. We remove the last category as it constitutes missing responses. This makes the response variable binary and hence a binary probit model is utilized to analyze the response based on the following set of covariates: age, income, household size, past use of marijuana, gender, education, state of residence, race, party affiliation and religion.

Age was recorded in years. Income (measured in US Dollars) was reported as belonging to one of 9 groups (0-10k, 10k-20k, $\cdots$, 40k-50k, 50k-75k, 75k-100k, 100k-150k, 150k and above, where 'k' denotes thousand). We convert income to a continuous variable by taking the mid-point of the

first 8 income groups and impute 150k for the last group. Household size represents the number of members in the family. Past use of marijuana and gender are indicator variables in the model. Educational attainment of the respondents is classified into three categories and the category 'high school & below' forms the base or reference category in the regressions. The variable 'tolerant states' indicates if a respondent lives in one of the 20 states, where recreational usage is legal, possession is decriminalized and/or allowed for medical use only.[7] Race is classified into three categories and 'White' race is used as the base category. The category 'Other Races' comprises of Asian, Hispanic, native American, Pacific Islanders and remaining races. Party affiliation is also classified into three categories and Republican Party is used as the reference category. Religion is classified into five categories and 'Protestant' is used as the base category. Here, the category 'Conservative' comprises of respondents belonging to one of the following religions: Buddhist, Hindu, Islam, Jew, Mormon and orthodox church. The category 'Liberal' comprises of respondents who claim to be Agnostic, Atheist, Universalist or nothing in particular. The descriptive statistics for all the variables are presented in Table 1.

Let us now look at the socio-demographic characteristics of a typical respondent in the sample. An average respondent is about 51 years old and s/he belongs to a household of size 3 with an annual income of 60,527 US Dollars. The sample is almost evenly split between males and females and 46.87 percent of respondents have a history of marijuana use. In the sample, the largest proportion of respondents (36.04 percent) have 'bachelors and above' degree, followed by 'below bachelors' degree (30.46 percent). A significant fraction of the respondents (31.64 percent) reside in states that have some favorable laws towards marijuana. The sample is predominantly White (79.36 percent) with a good representation (12.01 percent) of the African American population. Amongst the respondents, almost 30 percent consider themselves as Republican, about 34 percent declare themselves as Democrats and the remaining are Independent or belong to other parties. With respect to religious codification, the largest proportion (41.79 percent) is Protestants, followed by Roman Catholics (21.83 percent). A good proportion, 11.68 percent, declare themselves to be

[7]The list of tolerant states before the date of the survey include Alaska, Arizona, California, Colorado, Connecticut, Delaware, Hawaii, Maine, Maryland, Massachusetts, Michigan, Montana, Nevada, New Jersey, New Mexico, Oregon, Rhode Island, Vermont, Washington, Washington DC. Source: https://www.whitehouse.gov/ondcp/state-laws-related-to-marijuana.

simply Christian. The Liberal category forms about 18.61 percent and the Conservatives have the lowest fraction at 6.09 percent.

6.1.2. Estimation

We estimate four different binary probit models and present the estimated coefficients and standard errors for each model in Table 2. To begin with, Model 1 considers a basic set of covariates that includes log age, income, past use of marijuana, gender, education categories, household size and state of residence of the respondents. Subsequent models generalize Model 1 by adding more variables to the basic set of regressors. Specifically, Model 2 adds the race variable, Model 3 adds party affiliation to the list of variables in Model 2, and Model 4 incorporates religious denomination to the regressors in Model 3. A look at the model fitting measures in Table 2 reveal that all the four models have high LR statistic (see equation (12) in Section 4). Consequently, we reject the null hypothesis that the coefficients are jointly zero in each model. The two other goodness of fit statistics, McFadden's R^2 and hit-rate (defined in equations (13) and (14), respectively), also show that all models provide a good fit, with each subsequent model providing a better fit than the previous model. For the best fitting model i.e., Model 4, McFadden's R^2 is 0.15 and hit-rate is 69.03 percent. The latter implies that Model 4 correctly predicts more than 69 percent of the outcomes.

6.1.3. Results

We focus on the results from Model 4, since it provides the best model fit. Table 2 shows that log age has a negative coefficient and is statistically significant at 5 percent level.[8] This implies that young people are more supportive of marijuana legalization. This is not surprising since risk taking or deviant behavior is high amongst the younger population and is well documented in the literature (Brown et al, 1974). Moreover, Saieva (2008) reports a negative association between age and probability of supporting legalization, while Alfonso and Dunn (2007) and Delforterie et al (2015) note that marijuana prevalence rate is higher among the younger population.

The coefficient for income is negative, but statistically insignificant as also documented in Nielsen (2007) and Saieva (2008). As such, we do not find a significant relationship between income

[8]The default significance level is 5 percent and henceforth reference to significance level will be omitted.

Table 2: Estimation results for the binary probit model.

	MODEL 1		MODEL 2		MODEL 3		MODEL 4	
	COEF	SE	COEF	SE	COEF	SE	COEF	SE
INTERCEPT	1.74**	0.61	1.70**	0.64	0.11	0.66	0.51	0.69
LOG AGE	−0.41**	0.11	−0.41**	0.11	−0.39**	0.11	−0.29**	0.12
LOG INCOME	−0.06	0.05	−0.06	0.05	−0.02	0.05	−0.03	0.05
PAST USE	0.81**	0.08	0.82**	0.08	0.82**	0.08	0.81**	0.08
MALE	0.18**	0.08	0.17**	0.08	0.21**	0.08	0.17**	0.08
BACHELORS & ABOVE	0.23**	0.10	0.23**	0.10	0.23**	0.11	0.21*	0.11
BELOW BACHELORS	0.19*	0.10	0.19*	0.10	0.24**	0.10	0.26**	0.10
HOUSEHOLD SIZE	−0.04	0.03	−0.05	0.03	−0.04	0.03	−0.03	0.03
TOLERANT STATES	0.13	0.08	0.12	0.08	0.12	0.08	0.08	0.09
AFRICAN AMERICAN	..	..	−0.05	0.12	−0.26**	0.13	−0.14	0.13
OTHER RACES	..	..	0.12	0.14	0.03	0.14	0.02	0.15
DEMOCRAT	..	..	..	..	0.68**	0.10	0.57**	0.11
OTHER PARTIES	..	..	..	..	0.48**	0.10	0.40**	0.10
ROMAN CATHOLIC	..	..	..	..	..	..	0.18*	0.10
CHRISTIAN	..	..	..	..	..	..	−0.06	0.13
CONSERVATIVE	..	..	..	..		..	0.44**	0.18
LIBERAL	..	..	..	..	..	..	0.66**	0.12
LR (χ^2) STATISTIC		164.37		165.38		211.44		248.55
MCFADDEN'S R^2		0.10		0.10		0.13		0.15
HIT-RATE		66.67		66.58		67.51		69.03

** $p < 0.05$, * $p < 0.10$

and probability of supporting marijuana legalization. This is in disagreement with the hypothesis that economically weaker section will support legalization since marijuana is popular within the lower income group. Past use of marijuana may drive support for legalization, but this was not controlled either in Nielsen (2007) or Saieva (2008). Controlling for this variable in our models, we find that the coefficient for past use is largest amongst all variables and highly significant. This finding provides support to the hypothesis that individuals who have used marijuana in the past strongly favor its legalization and the large coefficient value implies that past use is an important factor in favoring legalization. The coefficient for male is positive and significant, indicating that males are more supportive of legalizing marijuana as compared to females, a result also supported by Nielsen (2007) and Delforterie et al (2015). Similarly, Rodríguez (2015) finds that boys are more likely to use marijuana during their adolescent years. Since past use is an important determinant

for supporting legalization and marijuana use is more prevalent among males, it is not surprising to find that males are more supportive of legalization.

The indicator variables for higher education i.e., 'bachelors & above' and 'below bachelors' have positive coefficients and are statistically significant (either at 10 or 5 percent significance level) relative to the base category, 'high school & below'. Thus, higher education leads to increased support for legalization possibly because a more educated individual can better understand the costs and medicinal benefits of marijuana. However, some studies have found that early use of marijuana leads to lower educational attainment and poor performance in school (Lynskey and Hall, 2000; Van Ours and Williams, 2009; Horwood et al, 2010). Household size has a negative effect, but the coefficient is not statistically significant. Similarly, the coefficient for 'tolerant states' is positive, but insignificant. This implies that residing in one of 20 states that offers some relaxation on marijuana use/offence does not statistically increase the probability of supporting legalization.

The coefficient for African American and 'Other Races', are not statistically different from the base category, White, with the exception of Model 3. The negative coefficient for African American, although insignificant, is rather surprising because one would expect African Americans to support legalization in order to curtail the large number of marijuana related arrests from the African American community. In line with this, Chen and Killeya-Jones (2006) also examine the extent of marijuana use across race and find that marijuana use is higher among suburban White students compared to their African American counterparts. Moreover, Nasim et al (2007) document the cultural orientation for African American young women and find that traditional religious beliefs and practices could be the reason behind less marijuana usage among African American.

Political affiliation often represents ideological differences towards any public policy and several poll studies conducted by Pew and Gallup have found that Republicans (Democrats) are more likely to oppose (favor) legalization of marijuana. We also arrive at a similar conclusion, with the coefficients for Democrat and 'Other Parties' being positive and significant. This suggests that individuals having either political orientations are more supportive of legalization compared to Republicans and the result is consistent with Nielsen (2007). Lastly, we look at the effect of religious affiliations since religious beliefs sometime acts as a protective factor against alcohol usage and smoking. The results show that the coefficient for Roman Catholic is positive and statistically significant at 10 percent, while coefficient for Christian is negative but statistically insignificant. In

contrast, the coefficients for Conservative and Liberal are both positive and statistically significant and hence both groups are more supportive of legalization compared to Protestants. However, individual opinions often do not strictly adhere to religious codes and conducts, so these results may vary significantly across samples.

The above discussion suggests that the signs of the coefficients, except the race variables, are consistent with what one would typically expect. However, the coefficients by themselves do not give the covariate effects (see Section 2 and 3). Table 3 presents the average covariate effects for all significant variables, either at 5 or 10 percent level. Results show that an increase in age by 10 years decreases the probability of support by 1.9 percent. The highest positive impact comes from past use, which shows that an individual who has used marijuana is 28.5 percent more likely to support legalization relative to someone who has never used it. Males are 5.7 percent more likely to support legalization relative to females. Higher education increases the probability of support and an individual with bachelors or higher degree (below bachelors) is 6.8 (8.5) percent more likely to support legalization relative to an individual with a high school degree or below. Political affiliation to the Democratic Party increases the probability of support by 18.8 percent. Similarly, an individual who identifies themselves with Independent and other parties is 13.4 percent more likely to support legalization compared to a Republican. Finally, an individual who is Conservative (Liberal) is 14.3 (22) percent more likely to support legalization relative to a Protestant.

Table 3: Average covariate effects from Model 4.

COVARIATE	ΔP(favor legalization)
AGE, 10 YEARS	-0.019
PAST USE	0.285
MALE	0.057
BACHELORS & ABOVE	0.068
BELOW BACHELORS	0.085
DEMOCRAT	0.188
OTHER PARTIES	0.134
ROMAN CATHOLIC	0.058
CONSERVATIVE	0.143
LIBERAL	0.220

6.2. Ordinal Probit Model - The Extent of Marijuana Legalization

The year 2013 was the first year in four decades that majority of Americans favored legalization of marijuana. While its support has grown overtime as shown in Figure 4, it is important to distinguish between different levels of support. This distinction is crucial because support for personal use of marijuana is stronger than support for its medical use and has different policy implications. Individuals may opine to support marijuana for medicinal benefits, but not for personal use. The February 2014 Political Survey recorded individual response as a three level categorical variable, which permits use of an ordinal probit model to study the effect of covariates on public opinion about the extent of legalization.

6.2.1. Data

The February 2014 Political Survey was conducted during February 14-23, 2014 by the Princeton Survey Research Associates and sponsored by the Pew Research Center for the People and the Press. In the survey, a representative sample of 1,821 adults living in the US were interviewed over telephone with 481 (1,340) individuals interviewed over land line (cell phone, including 786 individuals without a land line phone). The sampled data contain several missing observations and many respondents were unsure about their opinion on legalization. As before, we remove data on these respondents and are left with a sample of 1,492 observations for the study[9].

The dependent variable in the model is the respondents' answer to the question, "Which comes closer to your view about the use of marijuana by adults?". The options provided were, 'It should not be legal,' 'It should be legal only for medicinal use,' or 'It should be legal for personal use'. The fourth category labeled, 'Don't know/Refused' is removed from the study. Similar to the March 2013 Survey, the February 2014 Survey also collected information on the age, income, household size, past use, gender, education, race, party affiliation and religion. We use these variables as independent variables in the models. All the definitions and categories for the variables remain the same as in Section 6.1.1. We also include the indicator variable 'tolerant states', with the definition modified to include Illinois and New Hampshire to the previous list of 20 states.[10] Finally, we

[9]The data file ("Feb14Data.xlsx") and Matlab codes file ("OrdinalProbitScript.m") for estimation and post-estimation constructs are available at: `https://www.arshadrahman.com/Research.html`.

[10]Note that marijuana related laws were passed in Illinois and New Hampshire after the March 2013 Political Survey, but before the February 2014 Political Survey.

include an additional variable, labeled 'eventually legal', for which data was collected only in the February 2014 Survey. This variable indicates whether respondents expect marijuana to be legal irrespective of their individual opinion. We present the descriptive statistics for all the variables in Table 4.

Upon exploration of the socio-demographic characteristics of the current sample, we note that an average respondent is about 45.5 years old and s/he belongs to a household of size 3 with an annual income of 60,647 US Dollars. Thus, the typical respondent is about 6 years younger compared to the March 2013 data and approximately has the same household size and income. The percentage of males is higher in the current sample by 5 percent, but still close to a fair split between males and

Table 4: Descriptive summary of the variables (February 2014 Political Survey).

VARIABLE		MEAN	STD
LOG AGE		3.72	0.44
LOG INCOME		10.63	0.98
HOUSEHOLD SIZE		2.74	1.42
	CATEGORY	COUNTS	PERCENTAGE
PAST USE		719	48.19
MALE		792	53.02
EDUCATION	BACHELORS & ABOVE	551	36.93
	BELOW BACHELORS	434	29.09
	HIGH SCHOOL & BELOW	507	33.98
TOLERANT STATES		556	37.27
EVENTUALLY LEGAL		1,154	77.35
RACE	WHITE	1149	77.01
	AFRICAN AMERICAN	202	13.54
	OTHER RACES	141	9.45
PARTY AFFILIATION	REPUBLICAN	333	22.32
	DEMOCRAT	511	34.25
	INDEPENDENT & OTHERS	648	43.43
RELIGION	PROTESTANT	550	36.86
	ROMAN CATHOLIC	290	19.44
	CHRISTIAN	182	12.20
	CONSERVATIVE	122	8.18
	LIBERAL	348	23.32
PUBLIC OPINION	OPPOSE LEGALIZATION	218	14.61
	LEGAL ONLY FOR MEDICINAL USE	640	42.90
	LEGAL FOR PERSONAL USE	634	42.49

females. Similarly, the sample is almost evenly split between respondents who have used marijuana and those who have not. The largest proportion of respondents (36.93 percent) have 'bachelors & above' degree, followed by 'high school & below' degree (33.98 percent). A significant fraction of the respondents (37.27 percent) reside in states that have some favorable law on marijuana. Looking at the additional variable 'eventually legal', note that 77.35 percent of the surveyed people expect marijuana to be legal irrespective of their opinion. Similar to the earlier data, the sample is predominantly White (77.01 percent) with a good representation (13.54 percent) of the African American population. Party affiliation shows that 34.25 percent of the sample is comprised of Democrats, 22.32 percent Republicans and the remaining fraction are 'independent & others'. With respect to religious classifications, the largest proportion of respondents are Protestant (36.86 percent), followed by Liberal (23.32 percent) and Roman Catholics (19.44 percent).

6.2.2. Estimation

The ordinal probit model results are presented in Table 5, which show the coefficient estimates and standard errors of four different models. The estimation of models follows a similar sequence as in Table 2. Model 5 is the base model and contains log age, log income, past use of marijuana, male, education categories, household size, tolerant states and eventually legal. Model 6 adds the race categories to Model 5 and Model 7 adds party affiliation to the list of variables in Model 6. Finally, Model 8 contains all the variables in Model 7 and religious categories. The goodness of fit statistics are presented in the last three rows of Table 5. The LR statistics are large and each model fits better than the respective intercept model. The other two measures, McFadden's R^2 and hit-rate (defined in equations (13) and (14), respectively), show that Model 7 and Model 8 outperform the remaining two models. Model 8 provides a better fit compared to Model 7 as per McFadden's R^2 (0.1187 compared to 0.1254), whereas Model 7 has a marginally better hit-rate compared to Model 8 (59.11 percent as opposed 58.91 percent). So, the model correctly predicts approximately 60 percent of observed outcomes.

6.2.3. Results

We focus on the results from Model 8 because it is the most general model, provides the best fit according to McFadden's R^2 and no variables change sign. The results indicate that log age has a negative effect on the support for personal use (third category) and is statistically significant at 5

Table 5: Estimation results for the ordinal probit model.

	MODEL 5		MODEL 6		MODEL 7		MODEL 8	
	COEF	SE	COEF	SE	COEF	SE	COEF	SE
INTERCEPT	1.26**	0.44	1.27**	0.45	0.83*	0.46	0.34	0.48
LOG AGE	−0.44**	0.07	−0.45**	0.07	−0.45**	0.08	−0.35**	0.08
LOG INCOME	0.07*	0.03	0.07*	0.03	0.08**	0.03	0.09**	0.04
PAST USE	0.74**	0.06	0.73**	0.06	0.71**	0.06	0.69**	0.06
MALE	0.06	0.06	0.06	0.06	0.07	0.06	0.06	0.06
BACHELORS & ABOVE	0.26	0.08	0.26**	0.08	0.25**	0.08	0.24**	0.08
BELOW BACHELORS	0.06	0.08	0.05*	0.08	0.05	0.08	0.05	0.08
HOUSEHOLD SIZE	−0.04*	0.02	−0.04	0.02	−0.03	0.02	−0.02	0.02
TOLERANT STATES	0.11*	0.06	0.13**	0.06	0.09	0.06	0.07	0.07
EVENTUALLY LEGAL	0.58**	0.07	0.58**	0.07	0.56**	0.07	0.57**	0.07
AFRICAN AMERICAN	..	..	0.11	0.09	−0.01	0.10	0.03	0.10
OTHER RACES	..	..	−0.18*	0.11	−0.26**	0.11	−0.27**	0.11
DEMOCRAT	..	..	..	..	0.48**	0.09	0.44**	0.09
OTHER PARTIES	..	..	..	..	0.40**	0.08	0.36**	0.08
ROMAN CATHOLIC	..	..	..	..	..	..	0.10	0.09
CHRISTIAN	..	..	..	..	..	..	0.16	0.10
CONSERVATIVE	..	..	..	..		..	0.09	0.12
LIBERAL	..	..	..	..	..	..	0.39**	0.09
CUT-POINT	1.43**	0.05	1.43**	0.05	1.45**	0.05	1.46**	0.05
LR (χ^2) STATISTIC		316.27		321.47		356.99		377.02
MCFADDEN'S R^2		0.10		0.11		0.12		0.12
HIT-RATE		57.77		57.44		59.11		58.91

** p < 0.05, * p < 0.10

percent level. Alternatively, log age has a positive effect on opposing legalization (first category) and is statistically significant. However, the effect of age on medicinal use only (second category) cannot be determined *a priori*. Henceforth, we shall only discuss the effect on personal use and the impact on opposing legalization will be opposite to that of personal use. As before, the default level of significance used is 5 percent and further discussion will omit reference to significance level.

We note that the coefficient for log income is positive and statistically significant. This implies that individuals with higher income are more likely to support legalization for personal use. Past use of marijuana has a statistically significant positive effect on the probability of supporting personal use of marijuana. Moreover, the coefficient for past use is largest among all the variables, a result which is similar to that obtained in the binary probit model. Contrary to the finding in Section 6.1.3,

the coefficient for male is positive, but is not statistically significant. Thus, the current data do not confirm any role of gender on public opinion towards marijuana. Higher education is positively associated with support for personal use of marijuana. However, only the coefficient for 'bachelors degree & above' is statistically significant. The coefficients for household size and 'tolerant states' are not significant and conform to the results from the binary probit model. We find that 'eventually legal' has a significant positive effect, indicating that if an individual expects marijuana to be legal irrespective of his or her opinion, then s/he is more likely to support legalization for personal use.

The race variables suggest that opinions of African Americans on personal use of marijuana are not significantly different compared to the Whites. Such a similarity of opinions across race was also observed in the binary probit model. In contrast, 'Other Races' has a significant negative coefficient and is more opposed to legalization as compared to Whites. The coefficients for political party affiliations are in consonance with the results from Section 6.1.3. Affiliation to Democratic Party or 'Other Parties' increases the support for personal use and the coefficients are significant. Lastly, religious affiliations do not show a strong effect. Here, only the Liberals are more supportive of personal use of marijuana, while the opinions of the remaining religious categories are not significantly different from the base category, Protestant.

We mentioned in Section 3 that the coefficients of the ordinal probit model only give the direction of impact for the first and last categories, but not the remaining categories. The actual covariate effects need to be calculated for all the categories. We compute the average covariate effects for all significant variables and present them in Table 6. From the table, note that past use, eventually legal, and identifying oneself as a Democrat are three variables with the highest impact on public

Table 6: Average covariate effects from Model 8.

COVARIATE	ΔP(not legal)	ΔP(medicinal use)	ΔP(personal use)
AGE, 10 YEARS	0.015	0.012	−0.028
INCOME, $10,000	−0.005	−0.003	0.008
PAST USE	−0.129	−0.113	0.243
BACHELORS & ABOVE	−0.045	−0.035	0.080
EVENTUALLY LEGAL	−0.126	−0.060	0.186
OTHER RACES	0.059	0.031	−0.089
DEMOCRAT	−0.080	−0.066	0.147
OTHER PARTIES	−0.070	−0.051	0.121
LIBERAL	−0.068	−0.066	0.134

opinion. Past use of marijuana increases support for personal use by 24.3 percent and decreases the support for medicinal use and oppose legalization by 11.3 and 12.9 percents, respectively. Similarly, a respondent who expects marijuana to be legal is 18.6 percent more likely to support marijuana for personal use. This increase comes from a decrease in probability for medicinal use and oppose legalization, which are 6.0 and 12.6 percents, respectively. In the same way, a respondent who is a Democrat is 14.7 percent more likely to favor personal use, and 6.6 and 8.0 percents less likely to favor medicinal use and oppose legalization, respectively. The covariate effects for the remaining variables can be interpreted similarly.

7. Concluding Remarks

This chapter presents an overview of two discrete choice models, namely, binary and ordinal probit models which are prevalent in different disciplines including economics, finance, and sociology. We use the binary and ordinal probit models as prototypes to derive the likelihood and outline the estimation procedure, but the approach is general and applicable to its logit counterparts. Some interesting aspects about interpreting the coefficients of logit models are emphasized. Since discrete choice models are non-linear, the coefficients do not represent covariate effects. We explain how to compute the covariate effects with continuous and binary (indicator) variables. Three model fitting measures, namely, likelihood ratio statistic, McFadden's R-squared, and hit-rate are also described. After the theoretical discussion, the models are employed in two applications on marijuana legalization in the United States.

In the first application, a probit model is employed to analyze public response to legalization ('oppose legalization' or 'favor legalization') based on a host of individual and socio-economic variables. Data is taken from the March 2013 Political Survey collected by the Pew Research Center. The results show that age has a negative effect, but variables such as past marijuana use, male gender, and higher education have a positive effect on supporting legalization. Interestingly, past use of marijuana has the highest positive effect and increases the probability of support by 28.5 percent. In the second study, an ordinal probit model is employed to analyze ordered response to legalization ('oppose legalization', 'legal for medicinal use' or 'legal for personal use') using the February 2014 Political Survey data collected by the Pew Research Center. The results show that log age and 'Other Races' (non-White and non-African American) negatively (positively) affect the

probability of supporting personal use (oppose legalization) and the latter has the highest negative effect at 8.9 percent. Variables such as income, belief about legalization, and indicators for past use of marijuana, bachelors and above education, and affiliation to Democratic Party have a positive (negative) effect on personal use (oppose legalization). Amongst all variables, past use of marijuana has the highest positive effect on personal use at 24.3 percent.

The insights from these studies may assist policymakers to better assess public preference regarding marijuana legalization (particularly, medical marijuana) and the factors associated with such preferences. For instance, our finding that higher education increases support for legalization implies that providing information on the medicinal findings on marijuana and emphasizing college and university education is likely to increase support for it. Our findings may also be helpful to various advocacy groups engaged in promoting or opposing unrestricted legalization of marijuana. A clear understanding of the underlying factors that drive an individual's opinion will help these groups better plan their campaigns. For example, since support for legalization is negatively related with age, groups opposing legalization may consider not campaigning amongst the youth as it is unlikely to yield support. Similarly, lobbies engaged in opposing legalization may optimize their time use on better possibilities than convince an individual or group with a history of marijuana use to oppose legalization.

References

Abrams DI, Hilton JF, Leiser RJ, Shade SB, Elbeik TA, Aweeka FT, Benowitz NL, Bredt BM, Kosel B, Aberg JA, Deeks SG, Mitchell TF, Mulligan K, Bacchetti P, McCune JM, Schambelan M (2003) Short-term effects of cannabinoids in patients with HIV-1 infection: A randomized, placebo-controlled clinical trial. Annals of Internal Medicine 139(4):258–266

Abrams DI, Jay CA, Shade SB, Vizoso H, Reda H, Press S, Kelly ME, Rowbotham MC, Petersen KL (2007) Cannabis in painful HIV-associated sensory neuropathy. Neurology 68(7):515–521

Albert J, Chib S (1993) Bayesian analysis of binary and polychotomous response data. Journal of the American Statistical Association 88(422):669–679

Alfonso J, Dunn ME (2007) Differences in the marijuana expectancies of adoloscents in relation to marijuana use. Substance Use and Misuse 42(6):1009–1025

Benoit DF, Poel DVD (2010) Binary quantile regression: A Bayesian approach based on the asymmetric Laplace distribution. Journal of Applied Econometrics 27(7):1174–1188

Berman JS, Symonds C, Birch R (2004) Efficacy of two cannabis based medicinal extracts for relief of central neuropathic pain from brachial plexus avulsion: Results of a randomised controlled trial. Pain 112(3):299–306

Bresson G, Lacroix G, Rahman MA (2021) Bayesian panel quantile regression for binary outcomes with correlated random effects: An application on crime recidivism in Canada. Empirical Economics 60(1):227–259

Brown JW, Glaser D, Waxer E, Geis G (1974) Turning off: Cessation of marijuana use after college. Social Problems 21(4):527–538

Chen KW, Killeya-Jones LA (2006) Understanding differences in marijuana use among urban black and suburban white high school students from two U.S. community samples. Journal of Ethnicity in Substance Abuse 5(2):51–73

Delforterie MJ, Cremers HE, Agrawal A, Lynskey MT, Jak S, Huizink AC (2015) The influence

of age and gender on the likelihood of endorsing cannabis abuse/dependence criteria. Addictive Behaviors 42:172–175

Dimock M, Doherty C, Motel S (2013) Majority now supports legalizing marijuana, Pew Research Center. Available: `https://www.pewresearch.org/wp-content/uploads/sites/4/legacy-pdf/4-4-13-Marijuana-Release.pdf`, [Last accessed: 25 Jun 2023]

Drug Enforcement Administration (2011) Drugs of abuse. Tech. rep., Drug Enforcement Administration, U.S. Department of Justice, available: http://www.dea.gov/docs/drugs_of_abuse_2011.pdf

Duran M, Pérez E, Abanades S, Vidal X, Saura C, Majem M, Arriola E, Rabanal M, Pastor A, Farré M, Rams N, Laporte JR, Capellá D (2010) Preliminary efficacy and safety of an oromucosal standardized cannabis extract in chemotherapy-induced nausea and vomiting. British Journal of Clinical Pharmacology 70(5):656–663

Earleywine M (2005) Understanding Marijuana: A New Look at the Scientific Evidence. Oxford University Press, New York

Ellis RJ, Toperoff W, Vaida F, van den Brande G, Gonzales J, Gouaux B, Bentley H, Atkinson JH (2009) Smoked medicinal cannabis for neuropathic pain in HIV: A randomized, crossover clinial trial. Neuropsychopharmacology 34(3):672–680

Ferner M (2015) Obama: If enough states decriminalize marijuana, congress may change federal law. *The Huffington Post*, 17 March 2015, [Last accessed: 25 Jun 2023]

Geman S, Geman D (1984) Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence 6(6):721–741

Greenberg E (2012) Introduction to Bayesian Econometrics. 2nd Edition, Cambridge University Press, New York

Greene WH, Hensher DA (2010) Modeling Ordered Choices: A Primer. Cambridge University Press, Cambridge

Guzmán M (2003) Cannabinoids: Potential anticancer agents. Nature Reviews Cancer 3:745–755

Hastings WK (1970) Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57(1):97–109

Holmes CC, Held L (2006) Bayesian auxiliary variable models for binary and multinomial regression. Bayesian Analysis 1(1):145–168

Horwood LJ, Fergusson DM, Hayatbaksh MR, Najman JM, Coffey C, Patton GC, Silins E, Hutchinson DM (2010) Cannabis use and educational achievements: Findings from three Australasian cohort studies. Drug Alcohol and Dependence 110(3):247–253

Hosmer DW, Lemeshaw S, Sturdivant RX (2013) Applied Logistic Regression. 3rd Edition, John Wiley & Sons, Inc., Hoboken, New Jersey

Jeliazkov I, Rahman MA (2012) Binary and ordinal data analysis in economics: Modeling and estimation. In: Yang XS (ed) Mathematical Modeling with Multidisciplinary Applications, John Wiley & Sons Inc., New Jersey, pp 123–150

Jeliazkov I, Graves J, Kutzbach M (2008) Fitting and comparison of models for multivariate ordinal outcomes. Advances in Econometrics: Bayesian Econometrics 23:115–156

Johnson JR, Burnell-Nugent M, Lossignol D, Ganae-Motan ED, Potts R, Fallon MT (2010) Multicenter, double-blind, randomized, placebo-conrolled, parallel-group study of the efficacy, safety and tolerability of THC: CBD extract and THC extract in patients with intractable cancer-related pain. Journal of Pain and Symptom Management 39(2):167–179

Johnson VE, Albert JH (2000) Ordinal Data Modeling. Springer, New York

Kalant H (2004) Adverse effects of cannabis on health: An update of the literature since 1996. Progress in Neuro-Psychopharmacology 28:849–863

Koenker R, Bassett G (1978) Regression quantiles. Econometrica 46(1):33–50

Kordas G (2006) Smoothed binary regression quantiles. Journal of Applied Econometrics 21(3):387–407

Luce RD (1959) Indiviudal Choice Behavior: A Theoretical Analysis. John Wiley & Sons, New York

Luce RD, Suppes P (1965) Preferences, utility and subjective probability. In: Luce RD, Bush RR, Galanter E (eds) Handbook of Mathematical Psychology, vol 3, John Wiley & Sons Inc., New York, pp 249–410

Lynskey M, Hall W (2000) The effects of adolescent cannabis use on educational attainment: A review. Addiction 95(11):1621–1630

Maheshwari P, Rahman MA (2023) bqror: An R package for Bayesian quantile regression in ordinal models. The R Journal (forthcoming):1–17, URL `https://www.arshadrahman.com/Maheshwari_Rahman_bqror_Ver1.6.0.pdf`

Marschak J (1960) Binary choice constraints on random utility indications. In: Arrow KJ (ed) Stanford Symposium on Mathematical Methods in the Social Sciences, Stanford University Press, Stanford, CA, pp 312–326

McAllister SD, Murase R, Christian RT, Lau D, Zielinski AJ, Allison J, Almanza C, Pakdel A, Lee J, Limbad C, Liu Y, Debs RJ, Moore DH, Desprez P (2011) Pathways mediating the effects of cannabidiol on the reduction of breast cancer cell proliferation, invasion and metasis. Breast Cancer Research and Treatment 129(1):37–47

McCullagh P (1980) Regression models for ordinal data. Journal of the Royal Statistical Society – Series B 42(2):109–142

McFadden D (1974) Conditional logit analysis of qualitative choice behavior. In: Zarembka P (ed) Frontiers in Econometrics, Academic Press, New York, pp 105–142

McKelvey RD, Zavoina W (1975) A statistical model for the analysis of ordinal level dependent variables. The Journal of Mathematical Sociology 4(1):103–120

Meier MH, Caspi A, Ambler A, Harrington H, Houts R, Keefe RSE, McDonald K, Ward A, Poulton R, Moffitt TE (2012) Persistent cannabis users show neuropsychological decline from childhood to midlife. Proceedings of the National Academy of Sciences 109(40):2657–2664

Metropolis N, Rosenbluth AW, Rosenbluth M, Teller AH, Teller E (1953) Equation of state calculations by fast computing machines. Journal of Chemical Physics 21:1087–1092

Monroe AD (1998) Public opinion and public policy, 1980-1993. Public Opinion Quarterly 62(1):6–28

Nasim A, Corona R, Belgrave F, Utsey SO, Fallah N (2007) Cultural orientation as a protective factor against tobacco and marijuana smoking for African American young women. Journal of Youth and Adolescence 36(4):503–516

Nielsen AL (2007) Americans' attitudes toward drug-related issues from 1975-2006: The roles of period and cohort effects. Journal of Drug Issues 40(2):461–493

Ojha M, Rahman MA (2021) Do online courses provide an equal educational value compared to in-person classroom teaching? evidence from U.S. survey data using quantile regression. Education Policy Analysis Archives 29(85):1–25

Omata Y, Katayama H, Arimura TH (2017) Same concerns, same responses: A Bayesian quantile regression analysis of the determinants for nuclear power generation in Japan. Environmental Economics and Policy Studies 19(3):581–608

Paletz DL, Owen D, Cook T (2015) American Government and Politics in the Information Age. Flat World Knowledge

Poirier DJ (1995) Intermediate Statistics and Econometrics. The MIT Press, Cambridge

Polen MR, Sidney S, Tekawa IS, Sadler M, Friedman GD (1993) Health care use by frequent marijuana smokers who do not smoke tobacco. Western Journal of Medicine 158(6):596–601

Rahman MA (2016) Bayesian quantile regression for ordinal models. Bayesian Analysis 11(1):1–24

Rahman MA, Karnawat S (2019) Flexible Bayesian quantile regression in ordinal models. Advances in Econometrics 40B:211–251

Rahman MA, Vossmeyer A (2019) Estimation and applications of quantile regression for binary longitudinal data. Advances in Econometrics 40B:157–191

Rodríguez AMi (2015) Age, sex and personality in early cannabis use. European Psychiatry 30(4):469–473

Saieva AP (2008) Marijuana legalization: American's attitudes over four decades, M.S. Thesis

Shapiro RY (2011) Public opinion and American democracy. Public Opinion Quarterly 75(5):982–1017

Shepard EM, Blackley PR (2007) The impact of marijuana law enforcement in an economic model of crime. Journal of Drug Issues 37(2):403–424

Tanner MA, Wong WH (1987) The calculation of posterior distributions by data augmentation. Journal of the American Statistical Association 82(398):528–540

Train K (2009) Discrete Choice Methods with Simulation. Cambridge University Press, Cambridge

Van Ours JC, Williams J (2009) Why parents worry: Initiation into cannabis use by youth and their educational attainment. Journal of Health Economics 28(1):132–142

Wilsey B, Marcotte TD, Deutsch R, Gouaux B, Sakai S, Donaghe H (2013) Low dose vaporized cannabis significantly improves neuropathic pain. The Journal of Pain 14(2):136–148

Winterbourne M (2012) United States drug policy: The scientific, economic, and social issues surrounding marijuana. Available: http://web.stanford.edu/group/journal/cgi-bin/wordpress/wp-content/uploads/2012/09/Winterbourne_SocSci?_2012.pdf, [Last accessed 14 May 2015]

Yu K, Moyeed RA (2001) Bayesian quantile regression. Statistics and Probability Letters 54(4):437–447

Yu K, Zhang J (2005) A three parameter asymmetric Laplace distribution and its extensions. Communications in Statistics – Theory and Methods 34(9-10):1867–1879