Online Appendix to: Marijuana on Main Street? Estimating Demand in Markets with Limited Access

By Liana Jacobi and Michelle Sovinsky

This appendix provides details on the estimation methodology for various specifications and details of the counterfactual computations.
JEL: L15, K42, H2
Keywords: legalization, drugs, discrete choice models, limited information, selection correction

Estimation

A1. Model Fitting for Probit Model with Selection

For the estimation of the Probit model for marijuana use with selection based on binary access via MCMC methods we introduce the latent continuous access and marijuana use variables $\{a^*_{im}\}$ and $\{u^*_{im}\}$ and use the common latent variable representation of the probit

$$\begin{aligned}
a^*_{im} &= \mu^a_{im} + \eta_{im} = \tilde{\mathbf{h}}'_{im}\boldsymbol{\gamma} + \eta_{im}, \;\; a_{im} = I[a^*_{im} > 0] \\
u^*_{im} &= \mu^u_{im} + \varepsilon_{im} = \tilde{\mathbf{x}}'_{im}\boldsymbol{\beta} + \varepsilon_{im}, \;\; u_{im} = I[u^*_{im} > 0] \qquad if \quad a_{im} = 1
\end{aligned}$$

where for each sample subject $\tilde{\mathbf{h}}_{im}$ refers to the combined covariate vector for the access model containing intercept, individual attributes, state fixed effects, market-specific variables influencing access, and $\tilde{\mathbf{x}}_{im}$ is the combined covariate vector for the net utility model that contains the price p_{im}, individual attributes, market specific variables, year fixed effects, and state fixed effects in addition to the intercept. We define the vector of model parameters as $\boldsymbol{\theta} = (\boldsymbol{\gamma}, \boldsymbol{\beta}, \rho)$. Under the assumption that $(\eta_{im}, \varepsilon_{im}) \sim N_2(0, \Xi)$, where Ξ is $2x2$ covariance matrix with 1 on the diagonal and ρ on the off-diagonal and following to the definition of the likelihood contribution given in the paper equation 11, the likelihood of the model for all subjects in market m augmented with the latent access and net-use variables, $f(\mathbf{a}, \mathbf{u}, \{a^*_{im}\}, \{u^*_{im}\} | \boldsymbol{\theta}, \mathbf{W}, \{p_{im}\})$ can be expressed as

$$\begin{aligned}
&\prod_{i:a_{im}=0} \mathcal{N}(a^*_{im} | \tilde{\mathbf{h}}'_{im}\boldsymbol{\gamma}, \mathbf{1}) \; I[a^*_{im} \leq 0]^{1-a_{im}} \\
&\prod_{i:a_{im}=1} \mathcal{N}(u^*_{im} | \tilde{\mathbf{x}}'_{im}\boldsymbol{\beta} \;\; + \;\; \rho(a^*_{im} - \tilde{\mathbf{h}}'_{im}\boldsymbol{\gamma} \;, 1 - \rho^2) \\
&\quad \times \{ \; I[u^*_{im} \leq 0]^{1-u_{im}} + \; I[u^*_{im} > 0]^{u_{im}} \; \} \\
&\quad \times \mathcal{N}(a^*_{im} | \tilde{\mathbf{h}}'_{im}\boldsymbol{\gamma}, \mathbf{1}) I[a^*_{im} > 0]^{a_{im}}
\end{aligned}$$

where the inclusion of the latent data improves the tractability of the likelihood (Albert and Chib, 1993). The joint distribution of access and use for access subjects is now expressed in terms of the marginal-conditional decomposition. The indicator functions ensure that we choose the correct bivariate distribution with the distribution of latent use truncated according to the observed use.

For the Bayesian analysis we proceed with the common assumption of normal independent priors for the slope coefficients and correlation coefficient. The latter is restricted to the region $R = -1 < \rho < 1$ to ensure the positive definiteness of $\mathbf{\Xi}$. The joint prior is given by

$$\pi(\boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{\beta}|\mathbf{b}_0, \mathbf{B}_0)\ \mathcal{N}(\boldsymbol{\gamma}|\mathbf{g}_0, \mathbf{G}_0)\ \mathcal{N}(\rho|r_0, R_0) \times R$$

The prior means are set at zero. In combination with large prior variances this implies relatively uninformative prior assumptions. It should be noted that in the context of our very large data set the influence of the prior is very small as the information from the data via the likelihood will dominate the inference about the model parameters summarized in the posterior distribution. The posterior distribution, with the parameter space augmented by the latent access and marijuana variables, $\pi(\boldsymbol{\theta}, \mathbf{a}^*, \mathbf{u}^*|\mathbf{a}, \mathbf{u})$, is proportional to the product of the likelihood and the prior. We employ a straight forward Metropolis within Gibbs simulation algorithm with five blocks to generate draws from the posterior distribution of the parameter vector, as well as the marginal distributions of each parameter. By augmenting the parameter space with the latent access and net-use variables, the priors on the regression coefficients are conditionally conjugate, thus allowing for normal updates of slope the coefficients. The latent variables are also normal updates. A Metropolis Hastings update is used for the correlation parameter as the structure of the covariance matrix and the likelihood do not allow a Gibbs update. The detailed steps of the algorithm are as follows:

First, we draw a^*_{im} from $\mathcal{N}(a^*_{im}|\tilde{\mathbf{h}}'_{im}\boldsymbol{\gamma}, 1)\ I[a^*_{im} \leq 0]$ for $i \in \mathcal{I}_0$ and from $\mathcal{N}(a^*_{im}|\tilde{\mathbf{h}}'_{im}\boldsymbol{\gamma} + \rho(u^*_{im} - \widetilde{\mu}^u_{im}), 1-\rho^2)\ I[a^*_{im} > 0]$ for those subjects with $i \in \mathcal{I}$, where $i \in \mathcal{I}_0$ refers to the subset of subjects with no access and $i \in \mathcal{I}_1$ to those with access.

In the second step, we draw u^*_{im} for all subjects $i \in \mathcal{I}_1$ from either $\mathcal{N}(u^*_{im}|\tilde{\mathbf{x}}'_{im}\boldsymbol{\beta} + \rho(a^*_{im} - \tilde{\mathbf{h}}'_{im}\gamma), 1-\rho^2)\ I[u^*_{im} \leq 0]$ if $u_{im} = 0$ or from $\mathcal{N}(u^*_{im}|\tilde{\mathbf{x}}'_{im}\boldsymbol{\beta} + \rho(a^*_{im} - \tilde{\mathbf{h}}'_{im}\gamma), 1-\rho^2)\ I[u^*_{im} > 0]$ if $u_{im} = 1$.

In the third step, we draw $\boldsymbol{\gamma}$ from $\mathcal{N}(\hat{\boldsymbol{\gamma}}, \hat{\mathbf{G}})$ with

$$\hat{\boldsymbol{\gamma}} = \hat{\mathbf{G}}[\mathbf{G}_0^{-1}\mathbf{g}_0 + \sum_{i\in\mathcal{I}_0} \tilde{\mathbf{h}}_{im} a^*_{im} + \sum_{i\in\mathcal{I}_1} \tilde{\mathbf{h}}_{im}(1-\rho^2)^{-1}(a^*_{im} - \rho(u^*_{im} - \tilde{\mathbf{x}}'_{im}\boldsymbol{\beta})]$$
$$\hat{\mathbf{G}} = [\mathbf{G}_0^{-1} + \sum_{i\in\mathcal{I}_0} \tilde{\mathbf{h}}_{im}\tilde{\mathbf{h}}'_{im} + \sum_{i\in\mathcal{I}_1} \tilde{\mathbf{h}}_{im}(1-\rho^2)^{-1}\tilde{\mathbf{h}}'_{im}]^{-1}.$$

In the fourth step we draw $\boldsymbol{\beta}$ based on the subjects in $\mathcal{I}_1$ from $\mathcal{N}(\hat{\boldsymbol{\beta}}, \hat{\mathbf{B}})$ where

$$\hat{\boldsymbol{\beta}} = \hat{\mathbf{B}}[\mathbf{B}_0^{-1}\mathbf{b}_0 + \sum_{i\in\mathcal{I}_1} \tilde{\mathbf{x}}_{im}(1-\rho^2)^{-1}(u^*_{im} - \rho(a^*_{im} - \tilde{\mathbf{h}}'_{im}\boldsymbol{\gamma}))]$$
$$\hat{\mathbf{B}} = [\mathbf{B}_0^{-1} + \sum_{i\in\mathcal{I}_1} \tilde{\mathbf{x}}_{im}(1-\rho^2)^{-1}\tilde{\mathbf{x}}'_{im}]^{-1}.$$

In the last step we update ρ in Metropolis Hastings step based on the subjects in $\mathcal{I}_1$, since the conditional posterior distribution of ρ is not tractable. Following Chib and Greenberg (1998) we generate proposal value ρ' from a tailored student-t density $t_\nu(\mu, V)$ where μ is the mode of

$$\ln(\prod_{\mathcal{I}\in\mathcal{I}_1} \mathcal{N}(a^*_{im}, u^*_{im}|W_{im}\boldsymbol{\delta}, \Xi)\ , \text{ where}$$
$$\mathbf{W}_{im} = \begin{pmatrix} \tilde{\mathbf{h}}'_{im} \\ \tilde{\mathbf{x}}'_{im} \end{pmatrix},\ \boldsymbol{\delta} = \begin{pmatrix} \boldsymbol{\gamma} \\ \boldsymbol{\beta} \end{pmatrix} \text{ and } \Xi = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

and V is the inverse of the Hessian of the density evaluated at μ. The proposed value ρ' is accepted with probability

$$\alpha = \min\left(1\ ,\ \frac{\pi(\rho')\ \prod_{\mathcal{I}\in\mathcal{I}_1} \mathcal{N}(a^*_{im}, u^*_{im}|\mathbf{W}_{im}\boldsymbol{\delta}, \Xi')\ t_\nu(\rho|\mu, V)}{\pi(\rho)\ \prod_{\mathcal{I}\in\mathcal{I}_1} \mathcal{N}(a^*_{im}, u^*_{im}|\mathbf{W}_{im}\boldsymbol{\delta}, \Xi)\ t_\nu(\rho'|\mu, V)}\right).$$

We repeat the above steps for M iterations after an initial burn-in phase of M_0 iterations to allow for the convergence of the chain. We obtain a vector of M draws for each model parameter that reflects the (marginal) posterior distribution of each parameter. (Under the Bayesian approach the model parameters are random variables and all information about the parameters from the estimation is summarized in their respective posterior distributions.) In the main text we provide summaries of the posterior distributions in terms of the posterior means (coefficient estimate) and standard deviations or the 90% and 95% credibility intervals.

Simulated Prices

We address the issue of the unobserved individual prices as described in the paper in section III.A. by adding an additional step to the above described algorithm to draw the individual price for each subject i in market m with access from the market specific empirical price distribution (equation 7 in the paper). At the beginning of each iteration g of the algorithm we generate an individual price $p^{(g)}_{im}$ from

$$\sum_{t=1}^{3} \pi^g_{imt} \times p^g_{imt}\ , \quad \sum_{t=1}^{3} \pi^g_{imt} = 1$$

where t takes the values $t = 1, 2, 3$. The probability of using type t (π^g_{imt}) and the corresponding price for type t (p^g_{imt}) are drawn from the constructed empirical distributions (see equation 6) based on the observed data. Here we use normal distributions truncated at zero below and centered at the observed values of the prices and usage for each type in each market, $TN_{(0,\infty)}(\overline{p}_{mt}, \Omega^p_{mt})$ and $TN_{(0,\infty)}(\overline{\boldsymbol{\pi}}_{mt}, \Omega^\pi_{mt})$, respectively. The variances are based on observed variation in the data. By generating the price by type and usage of type from these distributions we can exploit the information on prices by type and market and usage of types within a market among users in the data, while at the same time allowing for some variation of prices and usage among subjects in a market. Note that different choices of distributions are possible and independence of the distributions across types is not necessary but chosen here based on the data restrictions. The graphs below show the resulting empirical price distributions for a subset of markets.

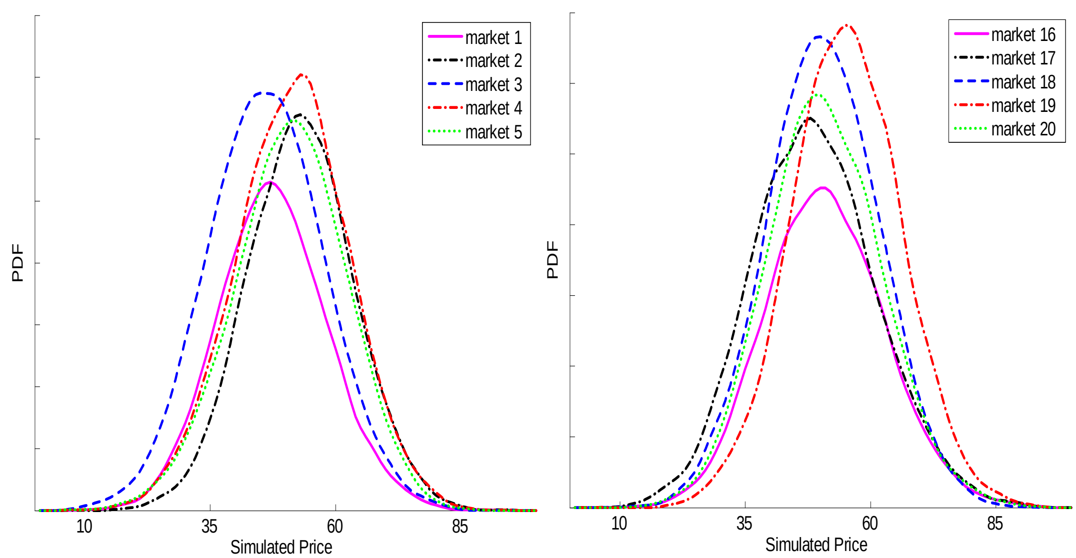

Figure A1. Simulated Price Distributions

A2. Marijuana Use Prediction Counterfactual

In the paper in Section V we report the probabilities of marijuana use for different counterfactual scenarios under various specifications of the selection model for the extensive margin of marijuana use. The reported probabilities are obtained using the standard Bayesian approach for prediction. The approach enables us to exploit all the information about the parameters summarized in the posterior distribution and to compute credibility intervals (Bayesian confidence intervals) for the probabilities of use and address the issue of unobserved individual prices via an empirical distribution generated as described above.

Let $n = 1$ refer to a random subject in market m from the sample, with demographic characteristics and market features in the use model and access model given by vectors $\tilde{\mathbf{x}}_{n+1,m}$, including price $p_{n+1,m}$, and $\tilde{\mathbf{h}}_{n+1,m}$, respectively. Under the selection model on marijuana use we can obtain the probability of marijuana use for the subject from the expression

$$\begin{aligned}\Pr(u_{n+1,m} &= 1|\mathbf{a},\mathbf{u}) = \\ \int \Pr(u_{n+1,m} &= 1|a_{n+1,m},\tilde{\mathbf{x}}_{n+1,m},\tilde{\mathbf{h}}_{n+1,m},p_{im}) \;\; d\widehat{P}_m(p_{n+1,m})\; dF_\pi(\boldsymbol{\theta}\;) \\ & \quad dF_{data}(\tilde{\mathbf{x}}_{n+1,m},\tilde{\mathbf{h}}_{n+1,m})\end{aligned}$$

where

$$\Pr(u_{n+1,m} = 1|a_{n+1,m},\tilde{\mathbf{x}}_{n+1,m},\tilde{\mathbf{h}}_{n+1,m},p_{im}) = \;\Phi(m_{n+1,m}|\boldsymbol{\theta},\tilde{\mathbf{x}}_{n+1,m},\tilde{\mathbf{h}}_{n+1,m},p_{im})$$

is the conditional probability of use (assuming access) with the conditional mean given by $m_{n+1,m} = \;\tilde{\mathbf{x}}'_{n+1,m}\,\boldsymbol{\beta} + \rho\eta_{n+1,m}$. The term $\rho\eta_{n+1,m}$ accounts for selection on unobservables, where the value of the unobservable term $\eta_{n+1,m}$ is found using information on the distribution of unobservables in the data by exploiting the data on the observed access of a subject within our model. Specifically, if $a_{n+1,m} = 0$, then it follows directly from the model of access that $\mu^a_{n+1,m} + \eta_{n+1,m} \leq 0$, or $\eta_{n+1,m} \leq -\mu^a_{n+1,m}$ where $\mu^a_{n+1,m} = \tilde{\mathbf{h}}'_{n+1,m}\boldsymbol{\gamma}$. As $\eta_{n+1,m}$ follows a standard normal distribution we generate $\eta_{n+1,m}$ from $TN_{(-\infty,-\mu^a_{n+1,m})}(0,1)$ for the subject. Similarly, for the case of $a_{n+1,m} = 1$ we have $\mu^a_{n+1,m} + \eta_{n+1,m} > 0$, so that we generate $\eta_{n+1,m}$ from $TN_{(-\mu^a_{n+1,m},\infty)}(0,1)$. Note that if we set $\rho\eta_{n+1,m} = 0$ we would implement the prediction based on the marginal model for marijuana use. While predictions are often based on the marginal model, this approach would lead us to considerably underpredict the benchmark case of use under pre-legalization relative to the observed use pre legalization, due to ignoring the important role that selection on unobservables plays in the context of marijuana use and more generally in the use of illicit drugs.

As indicated by the above integral expression, from the conditional probability we integrate out the prices based on the empirical price distribution $\widehat{P}_m(p_{im})$, the model parameters using the posterior distribution $F_\pi = \pi(\boldsymbol{\theta}\;|\mathbf{a},\mathbf{u})$ and the individual and market characteristics based on the empirical distribution of the sample. The above integral expression can be estimated in a straight forward manner using the draws from the posterior distribution from the MCMC algorithm discussed in the previous section. Essentially, at each iteration of the MCMC algorithm after the burn-in phase, we add an additional step where vectors $\tilde{\mathbf{x}}_{n+1,m}$ and $\tilde{\mathbf{h}}_{n+1,m}$ are drawn from the data and $\Phi(m_{n+1,m})$ is computed using the current MCMC draws on the model parameters and prices. The resulting vector of probabilities describes the predictive distribution of the probability of use. In tables 9 and 10 we report the mean probabilities and standard deviations of the predictive distributions.

We implemented the predictions under the following counterfactual scenarios s with the first scenario being the status quo:

Table A1—Counterfactual Scenarios

Scenario	Access	Legality	Price
1	no	no	pre-legality
2	yes	no	pre-legality
3	yes	yes	pre-legality
4	yes	yes	25% increase
5	yes	yes	cigarette price
6	yes	yes	price at marginal cost

For the prediction under these different scenarios let $\widetilde{v}_{n+1,m}$ denote the vector of market and demographic characteristics without the price and the disutility variables, $p_{n+1,m}$ and $L^{illegal}_{n+1,m}$. Let $\boldsymbol{\beta}_v, \beta_p$ and β_l denote the corresponding parameter (vectors) . We can then write the conditional probability of use under scenario s for our baseline model specification without interactions as

$$\begin{aligned}\Pr(u_{n+1,m} &= 1|a_{n+1,m}, \tilde{\mathbf{x}}_{n+1,m}, \tilde{\mathbf{h}}_{n+1,m}, p_{n+1,m}) \\ &= \Phi^{\,s}(\widetilde{v}'_{n+1,m}\,\boldsymbol{\beta}_v + p^s_{n+1,m}\beta_p + L^{illegal,s}_{n+1,m}\beta_l + \rho\eta_{n+1,m})\end{aligned}$$

where $p^s_{n+1,m}$ is drawn from an empirical distribution using the approach described above, with the mean of the price distributions for each type adjusted according to the assumed price change for scenarios 4 to 6. The disutility variable $L^{illegal,s}_{n+1,m}$ is set to zero for all subjects in scenarios that assume legality ($s > 2$), and otherwise remains as unchanged. Note that by predicting use for any random subject in the sample we assume that marijuana is accessible for all subjects. To account for limited access in our benchmark scenario $s = 1$ we follow the described approach, but set $\Phi(m_{n+1,m}) = 0$ if $a_{n+1,m} = 0$. For the predicted probabilities of use by various demographic groups presented in table 10, we draw $\widetilde{v}_{n+1,m}$ and $\tilde{\mathbf{h}}_{n+1,m}$ from the corresponding subsamples of subjects with no access and subjects with access. For those with no access we set $\Phi(m_{n+1,m}) = 0$ under scenario 1. Finally, the counterfactual use results in the paper in Section V are generated using our baseline model specification as well as our model specification with price and legality interaction terms. For the latter the expression in the mean probability of use $(p^s_{n+1,m}\beta_p + L^{illegal,s}_{n+1,m}\beta_l)$ is replaced with the corresponding interaction terms with age brackets.

A3. Model Fitting and Counterfactuals for Ordered Probit Model with Selection

We also estimate an ordered probit model with selection (intensive use margin model) for the discrete ordered marijuana frequency of use variable, $y_{im} = 0, 1, 2$, for the analysis of tax revenues in the paper in section V.C.:

$$\begin{aligned} y^*_{im} &= \tilde{\mathbf{x}}'_{im}\boldsymbol{\beta} + \nu_{im}, \text{ where } y_{im} = 0 \text{ if } y^*_{im} \leq 0, \; y_{im} = 1 \text{ if } 0 < y^*_{im} \leq \tau \\ \text{and } y_{im} &= 2 \text{ if } \tau < y^*_{im} \end{aligned}$$

where τ refers to the cut-off point that has to be estimated. The first cut-off point has been set to zero for identification purposes. The model for access remains unchanged and as before we assume that access and marijuana use may both be affected by unobserved factors so that $(\eta_{im}, \nu_{im}) \sim N_2(0, \Xi)$ where Ξ is $2x2$ covariance matrix with 1 on the diagonal and ρ on the off-diagonal. The likelihood of the model, $f(\mathbf{a}, \mathbf{y}, \{a^*_{im}\}, \{y^*_{im}\} | \boldsymbol{\xi}, \mathbf{W}, \{p_{im}\})$ where $\boldsymbol{\xi} = (\boldsymbol{\gamma}, \boldsymbol{\beta}, \rho, \tau)$, can be again expressed in terms of the latent data to improve the tractability of the likelihood (Albert and Chib, 1993) as

$$\begin{aligned} \prod_{i:a_{im}=0} \mathcal{N}(a^*_{im}|\tilde{\mathbf{h}}'_{im}\boldsymbol{\gamma}, \mathbf{1}) \; I[a^*_{im} &\leq 0]^{a_{im}} \\ \prod_{i:a_{im}=1} \mathcal{N}(a^*_{im}|\tilde{\mathbf{h}}'_{im}\boldsymbol{\gamma}, \mathbf{1}) I[a^*_{im} &> 0]^{1-a_{im}} \; \{\mathcal{N}(y^*_{im}|\tilde{\mathbf{x}}'_{im}\,\boldsymbol{\beta} + \rho(a^*_{im} - \tilde{\mathbf{h}}'_{im}\boldsymbol{\gamma}\;, 1-\rho^2) \\ \times(I[y_{im} &= 0] \; I[y^*_{im} \leq 0] + I[y_{im} = 1] \; I[0 < y^*_{im} \leq \tau] + \\ I[y_{im} &= 2] \; I[y^*_{im} > \tau])\} \end{aligned}$$

We again assume independent normal priors for $(\boldsymbol{\gamma}, \boldsymbol{\beta}, \rho)$ as in the probit model with selection (extensive use model). For the cut-off points it is sufficient to assume a priori that $\tau > 0$.

To simulate the posterior distribution $\pi(\boldsymbol{\xi}, \mathbf{a}^*, \mathbf{y}^* | \mathbf{a}, \mathbf{y})$ we employ a 6 step MCMC algorithm that is an extended and modified version of the 5 step algorithm for the Bivariate Probit with Selection discussed above. We add a 6th step to draw the cut-off point and also adjust the generation of the latent utility $\mathbf{y}^*$. For the latter, we draw y^*_{im} for all subjects $i \in I_1$ from the truncated normal distributions $\mathcal{TN}_{(a,b)}(y^*_{im}|\tilde{\mathbf{x}}'_{im}\boldsymbol{\beta} + \rho(a^*_{im} - \tilde{\mathbf{h}}'_{im}\boldsymbol{\gamma}), 1-\rho^2) \; I[a < y^*_{im} \leq b]$, where $(a = -\infty, b = 0)$ for $k = 0$, $(a = 0, b = \tau)$ for $k = 1$ and $(a = \tau, b = +\infty)$ for $k = 2$. To update the cut-off point we employ a Metropolis Hastings algorithm as the conditional posterior distribution is of an unknown form. To improve the performance we update the cut-off point marginalized over the latent utilities $\{y^*_{im}\}$ and generate the proposal values from the tailored student-t density $q(\tau) = t_{10}(\mu, V)$, where here μ is the mode of the likelihood of the access subjects with with y_{im}=1 and y_{im} =2, $f(\mathbf{a} = \mathbf{1}, \{a^*_{im}\}, \{y_{im}$=1$\}, \{y_{im}$=2$\} | \boldsymbol{\gamma}, \boldsymbol{\beta}, \rho, \mathbf{W})$ and V is the inverse of the Hessian of the density evaluated at μ. We maximize the proportional conditional likelihood expression (omitted $\mathcal{N}(a^*_{im}|\tilde{\mathbf{h}}'_{im}\boldsymbol{\gamma}, \mathbf{1}) I[a^*_{im} > 0]^{1-a_{im}}$ as it does not depend on cut-off points)

$$ln\left(\prod_{\mathcal{I}_1:\ y_{im}=1}\left[\Phi(\frac{\tau-m_{im}}{\sigma})-\Phi(\frac{-m_{im}}{\sigma})\right]\right)+ln\left(\prod_{\mathcal{I}_1:\ y_{im}=2}\left[1-\Phi(\frac{\tau-m_{im}}{\sigma})\right]\right)$$

where $m_{im} = \tilde{\mathbf{x}}'_{im}\boldsymbol{\beta} + \rho(a^*_{im} - \tilde{\mathbf{h}}'_{im}\boldsymbol{\gamma})$ and $\sigma = \sqrt{1-\rho^2}$. The maximization is subject to the constraint that $\tau > 0$.

The proposed value τ , with $\tau > 0$, is accepted with probability

$$\alpha = \min\left(1\ ,\ \frac{f(\{y_{im}{=}1\},\{y_{im}{=}2\}|\{a^*_{im}\},\mathbf{a}=\mathbf{1},\boldsymbol{\gamma},\boldsymbol{\beta},\rho,\tau',\mathbf{W})\ t_\nu(\tau|\mu,V)}{f(\{y_{im}{=}1\},\{y_{im}{=}2\}|\{a^*_{im}\},\mathbf{a}=\mathbf{1},\boldsymbol{\gamma},\boldsymbol{\beta},\rho,\tau,\mathbf{W})\ t_\nu(\tau'|\mu,V)}\right),$$

where again we use the conditional form of the likelihood of marijuana use, omitting the marginal likelihood of access as it does not depend on the cut-off point. As in the algorithm for the probit model we draw the prices for the access subjects from the corresponding empirical distribution at the beginning of each iteration of the algorithm. The estimates are presented in table 8.

For the estimation of the tax revenues in the paper in Section V.C. we again employ the Bayesian predictive approach described in Appendix A.A2. Instead of predicting the probability of use we predict the probability of use in each category k, $\widehat{G}_{ikr}$, for each subject i (in market m) in the sample under two different tax regimes $r = 1, 2$ from

$$\widehat{G}_{ikr} = \int \Pr(y_{im} = k|a_{i,m}, \tilde{\mathbf{x}}_{im}, \tilde{\mathbf{h}}_{im}, p^r_{im})\ \ d\widehat{P}_m(p^r_{im})\ dF_\pi(\boldsymbol{\xi}\ )\ \ dF_{data}(\tilde{\mathbf{x}}_{im}, \tilde{\mathbf{h}}_{im})$$

where as before we integrate over the price distribution, now also depending on the tax regime, the posterior distribution of posterior distribution of the relevant model parameters and the empirical distribution of the data (covariates). As in the extensive use model the prediction is based on the conditional probability of use, now for each category. For example, for $k = 0$ we have $\Pr(y_{im} = 0|\cdot) = \Phi(-\mu_{im})$ with $\mu_{im} = \tilde{\mathbf{x}}'_{im}\ \boldsymbol{\beta} + \rho\eta_{im}$ ($\tilde{\mathbf{x}}'_{im}\ \boldsymbol{\beta}$ includes age interactions). Under tax regime 1 with 25% tax on current prices, following equation 7 in the paper, the price the individual faces under the tax regime is generated from

$$p^1_{im} \sim \sum_{t=1}^{3} \int (\pi_{imt} * p^1_{imt})\ dF_\pi(\pi_{imt})\ dF_p(p^1_{mt})$$

where we adjust equation 6 in the paper and now have $p^1_{imt} \sim TN_{(0,\infty)}(1.25 * p_{imt}, \Omega^p_{mt})$ with the mean set at the current average market price of each type plus a 25% tax. Under tax regime 2 the price is set at the marginal cost of production of each type, MC_t, so that the price distribution simplifies to

$$p^2_{im} \sim \sum_{t=1}^{3} \int (\pi_{imt} * MC_t) \, dF_{\pi}(\pi_{imt})$$

As before the prices reflect a weighted average over the prices of three different types based on the usage probabilities of each type and its distribution $F_{\pi}(\pi_{imt})$.

Price Counterfactual Calculation

For the counterfactual price calculation under the ordered probit (intensive margin) with selection we find the counterfactual price that implies a predicted "post-legal" probability of no use among teenagers under legalization that is comparable to the probability of no use observed in the data before legalization. Given the model we implement the analysis at the market level, finding the counterfactual prices for all teenagers in market m to match the observed probability of no use in their market, call this S^{obs}_m. Let $S^{Post}_{im}(\widehat{\Theta}, data; p_{im})$ represent the probability of no use for teenager i in market m under a counterfactual of legalization, and p^{CF}_{im} the counterfactual price, where p^{CF}_{im} is chosen so that

$$\frac{\sum_{n_i=1}^{n_m} S^{Post}_{im}(\widehat{\Theta}, data; p_{im} = p^{CF}_{im})}{n_m} = S^{obs}_m$$

where n_m refers to the number of teenagers in market m and the estimated parameter set $\widehat{\Theta}$ (here means of the posterior distribution of the parameters). To find the counterfactual prices for each teenager in market m, we first predict post legalization probability of no use, as described in Appendix A.2, with price p_{im} coming from the corresponding (pre-legal) empirical distribution (equation 7 in the paper). We then find the counterfactual price, where p^{CF}_{im} is the price that equates

$$S^{Post}_{im}(\widehat{\Theta}, data; p^{CF}_{im}) = S^{obs}_m.$$

From our ordered probit model on marijuana use with selection it follows that the probability of no use is $S^{Post}_{im}(\widehat{\Theta}, data; p^{CF}_{im}) = \Phi(-(\mu_{im} = f(\widehat{\Theta}, data; p^{CF}_{im}))$, where μ_{im} is the conditional mean of marijuana use taking into account preferences and the selection of unobservables for teenager i so that under our model specification with price and age interactions

$$\Phi(-(\mu_{im} = p^{CF}_{im}\widehat{\beta}_{p,teen} + L^{illegal}_{im}\widehat{\beta}_l + \widetilde{v}'_{im}\, \widehat{\boldsymbol{\beta}}_v + \widehat{\rho}\eta_{im})) = S^{obs}_m,$$

where $L^{illegal}_{im}$ denotes the disutility from illegality variable, $\widetilde{v}_{n+1,m}$ the vector of demographic characteristics and market characteristics without the price. $\widehat{\beta}_{p,teen}$, $\widehat{\beta}_l$ and $\widehat{\boldsymbol{\beta}}_v$ and refer to the estimated coefficients (posterior means) on price for teenagers, the effect of disutility and the effects of the elements in $\widetilde{v}_{n+1,m}$, respectively. As before $\rho\eta_{im}$ accounts for selection on unobservables with $\eta_{im} \sim TN_{(-\infty,-\widehat{\mu}^a_{im},)}$ if $a_{im} = 0$ and $\eta_{im} \sim TN_{(-\widehat{\mu}^a_{im},\infty)}$ if $a_{im} = 1$ using the observed access before legalization (see also Appendix A.2). The counterfactual price is obtained from

$$p^{CF}_{im} = \max\left\{ \frac{-\Phi^{-1}(S^{obs}_m) - \widetilde{v}'_{im}\,\widehat{\boldsymbol{\beta}}_v - L^{illegal}_{im}\widehat{\beta}_l - \widehat{\rho}\eta_{im}}{\widehat{\beta}_{p,teen}}, 0 \right\},$$

where Φ^{-1} is the inverse of the normal CDF and the maximum condition ensures non-negative prices. The latter is needed as some teenagers have $S^{Post}_{im}(p_{im}) >= S^{obs}_m$ and for some teenagers their probability of no use is far above the market average at current prices and we would obtain a negative price to lower it to the market average level. (An alternative approach is to set $p^{CF}_{im} = p_{im}$ for teenager with $S^{Post}_{im}(p_{im}) >= S^{obs}_m$ and only compute the counterfactual price as described above for teenagers who's probability of no use is below the market average. This approach yields similar results.) In the main paper we provide summaries of the estimated counterfactual prices for teenagers needed to keep the proportion of non-users at pre-legal levels and also by gender. Since use remains illegal for teenagers the disutility variable $L^{illegal}_{im}$ remains unchanged.

*

REFERENCES

Albert, James H. and Siddhartha Chib. 1993. “Bayesian Analysis via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts.” *Journal of Business and Economic Statistics*, 11, 1-15.

Chib, Siddhartha and Edward Greenberg. 1998. “Analysis of Multivariate Probit Models.” *Biometrika*, 85, 347-361.