Decriminalization and Marijuana Smoking Prevalence: Evidence From Australia

Kannika Damrongplasit
Department of Health Services, University of California Los Angeles and Rand Corporation

Cheng Hsiao
Department of Economics, University of Southern California and Department of Economics and Finance, City University of Hong Kong and WISE, Xiamen University

Xueyan Zhao
Department of Econometrics and Business Statistics, and Centre for Health Economics, Monash University, Australia

January 16, 2009

Abstract

This paper uses the 2001 National Drug Strategy Household Survey to assess the impact of marijuana decriminalization policy on marijuana smoking prevalence in Australia. Both parametric and nonparametric methods are used. The parametric approach includes an endogenous probit switching, two-part, sample selection, and standard dummy variable models while the nonparametric approach uses propensity score stratification matching. Specification analyses are also conducted. We construct a nonparametric kernel-based test to select between parametric and nonparametric models, and use likelihood ratio test to choose among parametric models. Our analyses favor the endogenous switching model where decriminalization increases the probability of smoking by 16.2%.

Key words: Average Treatment Effect, Endogenous Probit Switching, Propensity Score Matching, Parametric and Nonparametric Specification Analysis, Bootstrapping
1. INTRODUCTION

Illicit drug usage is widespread around the world, posing significant social and economic costs to the health care, justice and social welfare systems in both developed and developing countries. According to the July 21, 2001 issue of the *Economist* magazine, the global retail sale of illegal drugs is estimated to be US$150 billion a year, which is in the same league as worldwide sale of tobacco and alcohol and about half the size of the pharmaceutical industry. Significant amounts of public funds have been spent by governments worldwide on dealing with the consequences of substance abuse and on educational programs. For example, the United States’ drugs policy costs approximately US$35 – US$40 billion a year while the Australian government has committed more than AUS$1 billion towards its National Illicit Drugs Strategy since 1997.

Among illicit drugs, marijuana is by far the most widely used. It is commonly considered a “softer” drug compared to other “harder” drugs such as cocaine, heroin, or amphetamines. The prevalence of hydroponic cultivation in recent years has also significantly improved the productivity of covert production. While there is more support for using marijuana for medical purposes in treating patients with nausea, glaucoma, spasm, and pain, much controversy has surrounded the detrimental health effects of recreational use of marijuana. Some suggest that marijuana use is linked to lung cancer, deteriorated immune systems, harmful effect on blood circulation, and short-term memory loss. For heavy users, there is also the problem of drug dependency and the related withdrawal symptoms such as anxiety and loss of appetite.

At the center of the controversy is whether legal sanction is the best approach to reduce the use and the associated harm of the drug. The ongoing debates of marijuana decriminalization concentrate on potential benefits and costs of such policy. A major supporting argument for
decriminalization is that a criminal charge is too severe a penalty relative to the crime itself. A criminal record can have many negative consequences on the subsequent life of an otherwise law-abiding person. For example, an offender may lose out in future employment opportunities or face problems in international travel. Furthermore, decriminalization would allow the government to separate the market of marijuana from the market of other harder drugs, thereby permitting the authorities to redirect their resources used in law enforcement and criminal justice system from the “softer” cannabis to “harder” drugs like cocaine, heroin, and amphetamines. For instance, a 2005 report by Jeffrey Miron entitled “The Budgetary Implications of Marijuana Prohibition” finds that legalizing marijuana can reduce the cost of enforcement in the United States by US$7.7 billion per year. Moreover, supporters also argue that, when marijuana is illegal, the young marijuana users are unnecessarily exposed to harder drug dealers at the same time, making it easier for them to move on to consume harder drugs. For those who argue against decriminalization, their first claim is that decriminalization inevitably lowers both the legal and social costs associated with the use of marijuana, thus sending a signal that it is acceptable to smoke marijuana. This may, in turn, encourage higher consumption of the drug as a result. Another reason against decriminalization is the gateway theory; that is, there is a growing concern that exposure to marijuana by youths may lead to their subsequent consumption of other harder drugs. Given the above debates, empirical evidence of the impact of marijuana decriminalization on marijuana usage is crucial. In particular, if decriminalization has little or no impact on smoking prevalence, then there is a strong argument in favor of the policy. On the other hand, if there is evidence that decriminalization significantly stimulates more marijuana smoking, then a liberal approach towards marijuana may not be as beneficial as advocated by its supporters.
Empirical results using various data sources from the United States are mixed. Saffer and Chaloupka (1995, 1998), and Pacula, Chriqui and King (2003) find the impact of decriminalization on marijuana smoking prevalence to be positive and significant. In contrast, DiNardo and Lemieux (2001), Pacula (1998), and Thies and Register (1993) discover insignificant effects of the marijuana policy reform on individual smoking decisions. There have been three empirical studies of the Australian experience. Cameron and Williams (2001), and Zhao and Harris (2004) have both found a positive and significant marginal effect of decriminalization on prevalence of about 2%, while Williams (2004) found it is only significant for the sub-sample of male aged 25 years old or above. Typically, binary probit models are used in these studies where the decriminalization dummy variable is treated as an exogenous explanatory variable.

This paper uses the 2001 Australian National Drug Strategy Household Surveys (NDSHS) to study the impact of marijuana decriminalization on marijuana usage. Australia consists of six states and two territories. As of 2001, South Australia, Australia Capital Territory, and Northern Territory had already decriminalized the possession and cultivation of small quantities of marijuana for personal consumption. Under this regulation, while supplying and cultivating commercial quantities of marijuana still attract severe criminal charges, an ‘on the spot’ fine has replaced the criminal charges for minor users. If one is caught of using or growing marijuana, one must pay a fine (usually between AUS$150-$200) within a specified period of time, usually within sixty days, in order to be eligible for the above reduced penalty of receiving no criminal record or imprisonment. If one fails to pay the fine, however, a criminal proceeding may follow possibly leading to a jail sentence. The policy of decriminalization is commonly known in Australia as “Cannabis Expiation Notice” system (CEN). Finally, for those states
which have not decriminalized marijuana, criminal offence for possessing, consuming or cultivating the drug is still retained.

This paper aims to assess the impact of decriminalization on marijuana smoking prevalence. There are three major differences between this paper and earlier studies. First, existing studies usually treat decriminalization as an exogenous dummy variable when performing regression analysis. This is also the case here because our data is based on the 2001 survey. As of 2001, three states in Australia had already decriminalized marijuana use namely, South Australia in 1987, Australia Capital Territory in 1992 and Northern Territory in 1996. Thus, in analyzing this set of data, we take state’s decriminalization decision as predetermined and focus on another source of joint dependence namely, marijuana smoking behavior and residential choice. At the individual level, the decision of living in a particular state may not be random and the decisions of where to live and whether to smoke may be related. Individuals may not be randomly selected to different states and there may be selection bias arising from those living in the decriminalized states versus those in the non-decriminalized states. This article attempts to address the potential endogeneity of marijuana smoking and the individual’s decision to reside in decriminalized versus non-decriminalized states by allowing individual’s marijuana smoking behavior equation to be correlated with his/her residential choice. Second, we provide a more flexible marijuana smoking behavior equation by allowing individuals to respond differently when the legal and institutional arrangement changes. Third, both parametric and nonparametric analyses are conducted and their reliabilities are examined.

Essentially, the advantages of parametric approach are the disadvantages of nonparametric approach and the advantages of nonparametric approach are the disadvantages of parametric approach. The advantages of parametric approach are that it can simultaneously take
account of selection on observables and unobservables (provided the model is correctly specified) and allows the (efficient) estimation of the effects of individual factors on the outcomes. The disadvantages of the parametric approach are that both the conditional mean functions of observable factors and the probability distributions of the effects of unobservable factors need to be specified. The advantages of nonparametric approach are that neither the conditional mean functions of observable factors nor the probability distributions of the effects of unobservable factors need to be specified. The disadvantages (of the propensity score matching) are that it only takes account of selection on observables and only estimates the treatment effects. We will discuss the pros and cons through our specification analyses.

The remainder of this paper is organized as follows. Section 2 presents an endogenous probit switching model as the maintained hypothesis and treat the traditional dummy variable approach, sample selection model and two-part model as its nested alternatives, respectively called binary probit, bivariate probit and two-part models. Section 3 provides a description of our data. Estimation results and the comparison of our findings to existing literature are reported in section 4. In section 5, we present alternative measures of treatment effect from the propensity score stratification matching method. Specification analyses are provided in section 6. Conclusions are in section 7.

2. THE MODEL

We assume that the utility for an individual’s residential choice ($d=1$ if residing in decriminalized state and 0 if not) and marijuana consumption ($M$) is separable from the utility of consuming other goods. Similar to Carneiro, Hansen and Heckman (2003), Keane and Wolpin (1997), etc., we assume that an individual’s utility function for marijuana consumption and residential choice is state dependent on the initial endowment and institutional arrangement of
the two residential regimes. In particular, we let an individual’s utility function be
\[ U(M, d \mid a, a^*) = d U_1(M \mid a) + (1 - d) U_0(M \mid a) + h(d \mid a^*) , \]
where \( U_1(\cdot) \) and \( U_0(\cdot) \) are the utility functions of consuming marijuana for an individual residing in decriminalized and non-decriminalized states, respectively; \( h(\cdot) \) denotes the utility of living in decriminalized state or non-decriminalized state; \( a \) and \( a^* \) are socio-demographic, institutional and idiosyncratic components that affect \( M \) and \( d \), respectively; \( a \) and \( a^* \) may contain overlapping elements. \( U_1(\cdot) \) and \( U_0(\cdot) \) are assumed to be different because the same amount of marijuana consumption may lead to different level of utility in decriminalized and non-decriminalized state due to different institutional setup in these two locations, for example, the risk of smoking marijuana could be different.

Maximizing utility subject to budget constraint \( I \) yields \( M_1(p, I \mid a) \geq 0 \) if an individual resides in a decriminalized state and \( M_0(p, I \mid a) \geq 0 \) if resides in a non-decriminalized state, where \( p \) denotes price of marijuana. Substituting \( M_1(\cdot) \) and \( M_0(\cdot) \) into the utility function yields the conditional indirect utility \( V^1(p, I \mid a, a^*) \) and \( V^0(p, I \mid a, a^*) \) for residing in decriminalized and non-decriminalized state, respectively. Then \( d = 1 \) if \( d^* = V^1(p, I \mid a, a^*) - V^0(p, I \mid a, a^*) > 0 \) and \( d = 0 \) if \( d^* \leq 0 \). Approximating \( M_1(\cdot) \) and \( M_0(\cdot) \) by
\[ y^*_1 = \alpha_1 + \beta_1 x + \epsilon_1, \quad (2.1) \]
\[ y^*_0 = \alpha_0 + \beta_0 x + \epsilon_0, \quad (2.2) \]
such that
\[ M_1(x) = y^*_1 \text{ if } y^*_1 > 0, \text{ and } M_1(x) = 0 \text{ if } y^*_1 \leq 0, \quad (2.3) \]
and
\[ M_0(x) = y^*_0 \text{ if } y^*_0 > 0, \text{ and } M_0(x) = 0 \text{ if } y^*_0 \leq 0, \quad (2.4) \]
where $x$ denotes the observable factors of $p$, $I$, and $a$ that affect the demand for marijuana and $\varepsilon_1$, $\varepsilon_0$ denote the effects of unobservable. Equation (2.1)-(2.4) imply \( \text{Prob}(M_1(x) = 0) = \int_{-\infty}^{\infty} f(\varepsilon_1) d\varepsilon_1 \) and \( \text{Prob}(M_0(x) = 0) = \text{Prob}(y^*_1 < 0 \mid x) = \int_{-\infty}^{\infty} f(\varepsilon_0) d\varepsilon_0 \).

We approximate $d^*$ by a reduced form specification

$$d^* = \gamma_1 x + \gamma_2 z + \nu,$$

such that $d = 1$ if $d^* > 0$ and 0 otherwise, where $z$ and $\nu$ denote the additional observable factors in $a^*$ that are not in $a$ and the effects of unobservable factors in both $a$ and $a^*$ that affect residential choice.

Our data are in the form $\phi(\gamma, d)$, where $y = 1$ indicates an individual is a marijuana smoker and 0 otherwise, and $d = 1$ if an individual resides in a decriminalized state and 0 otherwise. From (2.1)-(2.4), it follows

$$y = 1 \text{ if } d^* y^*_1 + (1 - d) y^*_0 > 0, \text{ and } y = 0 \text{ otherwise.} \quad (2.6)$$

Equations (2.1)-(2.6) lead to an endogenous switching model. The model is in a limited information framework where we have structural form specification for the demand for marijuana for individuals residing in decriminalized and non-decriminalized state, respectively, and a reduced form specification for the residential choice. The identification of the structural marijuana use equation is achieved through the excluded variables, $z$, that are important in predicting the residential choice (e.g. Hsiao (1983)).

Many of the conventional models become a special case of this model. For instance, we could have (a) $\rho_{1\nu} = \rho_{0\nu} \neq 0$, which is referred to as a restricted switching model, where $\rho_{1\nu}$ and $\rho_{0\nu}$ denote the correlations between $\varepsilon_1$ and $\nu$ and $\varepsilon_0$ and $\nu$, respectively. (b) When
\( \rho_{10} = \rho_{00} = 0 \), the residential choice equation’s error term is uncorrelated with the smoking equation. In this case, model (2.1)-(2.6) is a generalization to the frequently used two-part model (Duan, Manning, Morris and Newhouse 1983, 1984) or Quandt (1972) exogenous regime switching model. (c) When \( \beta_1 = \beta_0 = \beta \), and \( \rho_{10} = \rho_{00} \neq 0 \), the model is analogous to the sample selection model (Amemiya 1985) in which

\[
M = \alpha + \beta x + \delta d + \varepsilon \quad \text{if} \quad \alpha + \beta x + \delta d + \varepsilon > 0, \quad \text{and} \quad M = 0 \quad \text{otherwise,} \tag{2.7}
\]

where \( \alpha = \alpha_0 \) and \( \delta = \alpha_1 - \alpha_0 \). (d) When \( \rho_{10} = \rho_{00} = 0 \) and \( \beta_1 = \beta_0 \), model (2.1)-(2.6) reduces to dummy variable approach to evaluation the effect of decriminalization with \( M \) given by (2.7) and uncorrelated with (2.5).

If \( \alpha_1, \beta_1, \alpha_0, \beta_0 \) and density function of \( \varepsilon_1 \) and \( \varepsilon_0 \) are known, \( \text{Prob}(y = 1 \mid x) \) is known for parametric analysis. We assume \( (\varepsilon_1, \varepsilon_0, \upsilon) \) are jointly normally distributed with mean zeros and covariance \( V(\varepsilon_i) = V(\varepsilon_0) = V(\upsilon_i) = 1, \text{Cov}(\varepsilon_i, \varepsilon_0) = \rho_{10}, \text{Cov}(\varepsilon_i, \upsilon_i) = \rho_{10}, \text{and} \text{Cov}(\varepsilon_0, \upsilon_i) = \rho_{00} \). Therefore, we may estimate average treatment effect (ATE) by

\[
\int [\Phi(\alpha_1 + \beta_1 x) - \Phi(\alpha_0 + \beta_0 x)] f(x) dx \tag{2.8}
\]

and average treatment effect of the treated (ATET) by

\[
\int [\Phi(\alpha_1 + \beta_1 x) - \Phi(\alpha_0 + \beta_0 x)] f(x \mid d = 1) dx \tag{2.9}
\]

or their restricted version, where \( \Phi(a) \) denotes the integrated standard normal from \(-\infty \) to \( a \). If the sample is randomly drawn, the ATE or ATET may be approximated by

\[
\frac{1}{n} \sum_{i=1}^{n} [\Phi(\alpha_1 + \beta_1 x_i) - \Phi(\alpha_0 + \beta_0 x_i)], \tag{2.10}
\]

where the summation is over the complete sample or those residing in decriminalized state.
In what follows, we shall first present parametric analysis, then nonparametric analysis. We shall show that the results are sensitive to model specifications and shall discuss which model is likely to yield more accurate measurement of the effects of decriminalization.

3. DESCRIPTION OF VARIABLES AND DATA

3.1 Selection of Explanatory Variables

Marijuana is an addictive recreational drug, and studies on recreational drugs have arisen from many disciplines such as psychology, medicine, epidemiology and sociology, as well as economics. Maximizing the state dependent utility function leads to demand for marijuana as a function of price, income, as well as some standard socioeconomic and demographic variables that capture heterogeneity in demand such as age, gender, marital status, educational attainment, work status and ethnic background for the Australian indigenous population as commonly postulated in empirical studies (e.g. Becker and Murphy (1988), Pacula (1998), Williams (2004) and Zhao and Harris (2004)). The impact of the legal risk of smoking via the decriminalization policy is captured by allowing the responses to the conditional variables to be different for the two types of states using the endogenous switching model described above.

For residential choice, the existing literature (Feridhanusetyawan and Kilkenny 1996; Kittiprapas and McCann 1999) suggests that both personal and location characteristics are key determinants of individual’s residential choice. Since we are taking a limited information approach, in addition to those variables determining marijuana smoking, additional variables like number of dependent children, which might influence one’s residential choice, and unemployment rate in each individual’s state of residence that is a proxy for state-specific effects are also used to predict the residential choice as well as to provide exclusion restrictions needed to identify the marijuana use behavioral equations (e.g. see Hsiao (1983)).
3.2 Data

The data used come from three different sources: 2001 Australia National Drug Strategy Household Survey (NDSHS), Australia Bureau of Statistics (ABS), and Australian Illicit Drug Report. The NDSHS is a nationally representative household survey of non-institutionalized civilian Australian population with age 14 and older. It provides information on individual drug usage, and many socioeconomic and demographic variables. Three different survey methods were implemented: a drop-and-collect questionnaire, a face-to-face personal interview, and a computer assisted telephone interview. For more sensitive questions like individual drug usage, measures were put in place so that the information is kept confidential from the interviewer in order to minimize potential underreporting of drug use. There are altogether 26744 observations available in the 2001 wave of NDSHS. In addition, the 2001 NDSHS also provides information on explanatory variables such as household income ($Income$), age ($Age1419$, $Age2024$, $Age2529$, $Age3034$, $Age3539$, $Age4069$, $Age70$), gender ($Male$), marital status ($Married$, $Divorce$, $Widow$, $Never Married$), number of dependent children ($#~Depchild$), educational attainment ($Degree$), employment status ($Working status$), and ethnicity ($Aboriginal$). A dichotomous variable ($Decrim$) is also defined indicating whether a person resides in a decriminalized state. After deleting observations with missing data, the resulting sample is 14008. South Australia, Australia Capital Territory, and Northern Territory had already adopted decriminalization of marijuana by 2001, so observations from these three states are classified as the treatment group. Definition of all variables are listed in Table 1.

The price of marijuana is obtained from the Australia Bureau of Criminal Intelligence (ABCI 2002) and the Australian Crime Commissions (ACC 2003). Four different prices by state are available: (i) price of head per ounce, (ii) price of head per gram, (iii) price of leaf per ounce,
and (iv) price of leaf per gram. These prices are first converted to the same unit, price per ounce. Then, for each state, a weighted average of (i), (ii), (iii), and (iv) is computed by using proportions of the respondents’ form of purchase as weights. We also deflate each state’s weighted average price of marijuana by the state’s CPI, and apply logarithmic function. The final price of marijuana is denoted $P_{MAR}$. A thorough discussion on Australia’s marijuana price can be found in Clements (2004). State-level CPIs and state-level unemployment rates are drawn from the Australia Bureau of Statistics (ABS 2003a,b), where the later has its unit expressed in term of percentage (%). Table 1 provides summary statistics of dependent and independent variables for all observations, treatment observations and control observations.

**Table 1: Summary Statistics of Dependent and Independent Variables (N = 14008)**

<table>
<thead>
<tr>
<th>Variable and Definition</th>
<th>All Data (N = 14008)</th>
<th>Treatment (N = 2968)</th>
<th>Control (N = 11040)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
<td>Mean</td>
</tr>
<tr>
<td>y</td>
<td>0.157</td>
<td>0.364</td>
<td>0.181</td>
</tr>
<tr>
<td>Decrim</td>
<td>0.212</td>
<td>0.409</td>
<td>1</td>
</tr>
<tr>
<td>$P_{MAR}$ (log of real price of marijuana)</td>
<td>5.876</td>
<td>0.237</td>
<td>5.929</td>
</tr>
<tr>
<td>Income (log of real household annual income)</td>
<td>10.445</td>
<td>0.757</td>
<td>10.554</td>
</tr>
<tr>
<td>Age1419 (1 if age is 14-19, 0 if not)</td>
<td>0.054</td>
<td>0.227</td>
<td>0.047</td>
</tr>
<tr>
<td>Age2024 (1 if age is 20-24, 0 if not)</td>
<td>0.077</td>
<td>0.267</td>
<td>0.075</td>
</tr>
<tr>
<td>Age2529 (1 if age is 25-29, 0 if not)</td>
<td>0.100</td>
<td>0.300</td>
<td>0.103</td>
</tr>
<tr>
<td>Age3034 (1 if age is 30-34, 0 if not)</td>
<td>0.122</td>
<td>0.327</td>
<td>0.121</td>
</tr>
<tr>
<td>Age3539 (1 if age is 35-39, 0 if not)</td>
<td>0.126</td>
<td>0.331</td>
<td>0.130</td>
</tr>
<tr>
<td>Age4069 (1 if age is 40-69, 0 if not)</td>
<td>0.455</td>
<td>0.498</td>
<td>0.466</td>
</tr>
<tr>
<td>Male (1 if male, 0 if female)</td>
<td>0.476</td>
<td>0.499</td>
<td>0.488</td>
</tr>
<tr>
<td>Married (1 if married, 0 if not)</td>
<td>0.619</td>
<td>0.486</td>
<td>0.621</td>
</tr>
<tr>
<td>Divorce (1 if divorced, 0 if not)</td>
<td>0.115</td>
<td>0.319</td>
<td>0.124</td>
</tr>
<tr>
<td>Widow (1 if widowed, 0 if not)</td>
<td>0.038</td>
<td>0.191</td>
<td>0.033</td>
</tr>
<tr>
<td># Depchild (# of dependent children under 14)</td>
<td>0.594</td>
<td>0.895</td>
<td>0.583</td>
</tr>
<tr>
<td>Degree (1 if university degree, 0 if not)</td>
<td>0.260</td>
<td>0.438</td>
<td>0.292</td>
</tr>
<tr>
<td>Working Status (1 if unemployed, 0 if not)</td>
<td>0.028</td>
<td>0.164</td>
<td>0.019</td>
</tr>
<tr>
<td>Aboriginal (1 if Aboriginal, 0 if not)</td>
<td>0.013</td>
<td>0.111</td>
<td>0.017</td>
</tr>
<tr>
<td>State Unemployment rate (rate in %)</td>
<td>6.986</td>
<td>1.375</td>
<td>6.331</td>
</tr>
</tbody>
</table>
4. EMPIRICAL RESULTS

Maximum likelihood method is used to derive our parameter estimates. Because we do not simultaneously observe \( y_1^* \) and \( y_0^* \), the joint distribution of \((\varepsilon_1, \varepsilon_0)\) or \( \rho_{10} \) is not identified. Table 2 reports estimated coefficients for the marijuana use and the reduced form residential choice equations and the average treatment effects for all five models. Estimates of the marginal effects on marijuana participation probability are provided in Table 3 for a reference individual who is male, aged 14-19 years old, with less than university education, never married, not unemployed, not of Aboriginal origin, residing in a non-decriminalized state, having income equal to the mean of household income, and facing marijuana price equal to the mean price.

Our results show that decriminalization has positive and generally significant impacts on marijuana smoking behavior although their magnitudes differ across different models. We present the average treatment effect (ATE) of decriminalization for all models at the bottom of Table 2, and the marginal effects of specific factors for our reference person in Table 3. We first discuss each model’s estimated ATE here. When using a simple binary probit model without accounting for endogeneity of treatment and flexibility in behavior, ATE is estimated to be 3.7%. This is very similar to those estimated by Cameron and Williams (2001), and Zhao and Harris (2004). When accounting for endogeneity of treatment as in model (iv), ATE rises to 4%.

However, when allowing for behavioral differences between the treatment and the control groups but ignoring endogenous treatment as in the two-part model, we obtain ATE of 13.7%. Finally, in the endogenous probit switching models (i) and (ii), our estimated ATEs are found to be 16.2% and 18.3%, respectively. It is clear that the two-part and the endogenous probit switching models provide stronger support to the opponents of marijuana decriminalization policy because they yield substantially larger ATEs than the binary probit and the bivariate probit models.
For the impact of specific explanatory variables on marijuana use behavior, models (i), (ii) and (iii) yield results that are similar while models (iv) and (v) generate comparable outcomes. Consistent with a priori conjecture that demand is negatively associated with the price of marijuana, the coefficients of $P_{MAR}$ are negative and significant for all models, implying that there is negative own price responsiveness. For a reference person, a 10% increase in the price of marijuana decreases the probability of using it by 1.27% according to our binary probit and the bivariate probit models. When allowing for behavioral differences between the treatment and the control groups and/or taking into account possible endogeneity, we find the negative own price effect to be much larger in the decriminalized states. In particular, for a reference person, as marijuana price increases by 10%, the probability of using the drug is estimated to fall by approximately 21% in decriminalized states but only drop by 1.18% to 1.45% in non-decriminalized states. The higher price responsiveness in the decriminalized states is expected, as we would anticipate that price plays a much smaller role in the non-decriminalized states where the risk premium of being caught is expected to play a bigger role than price in the smoking decision.

If marijuana is a normal good, we would expect the coefficient of household income to be positive and significant. However, the coefficient of household income is negative but insignificant for all models. It indicates that income effect is absent in this study. Existing literature also reports mixed results on income. For example, the full-sample estimation by Saffer and Chaloupka (1998) finds that income has insignificant effect on the probability of marijuana use while Pacula (1998), and Thies and Register (1993) both report significant negative income effect.
As for other conditional variables that affect marijuana use behavior, age is an important factor affecting marijuana smoking behavior. Cameron and Williams (2001) and Williams (2004) both report that the probability of participating in marijuana peaks for people in the 20-24 years old age-group, and then monotonically declines for subsequent ages. Our maximum likelihood coefficients in Table 2 obtain this same finding for the treatment group of models (i), (ii) and (iii). However, for the rest of the models (i.e. the control group of models (i), (ii), and (iii), model (iv), and model (v)), the peak of smoking prevalence is found to occur among those aged 25-29 years old. Our results indicate that young adults have the highest risk of becoming marijuana smokers in decriminalized states while adults have the greatest exposure in non-decriminalized states.

We find positive and significant coefficient on the gender dummy variable across models just like previous studies. Married individuals are less likely to use marijuana compared to their never married counterparts. We also find that widowed individuals are less likely to be marijuana smokers for all models and for both the treatment and control groups. We find no difference in cannabis usage between those who are divorced and those who are single across all models.

Education attainment does not seem to play a role in the decision to consume marijuana when using models (iv) and (v). However, when allowing for behavioral differences due to policy change as in model (iii) and possible endogeneity of treatment as in models (i) and (ii), we find that the effects of education do differ across the treatment and the control groups. For those who live in decriminalized states, having a university degree substantially reduces their likelihood of becoming marijuana smokers. On the other hand, for those living in non-decriminalized states, there is no difference in marijuana smoking prevalence between those with and without tertiary education.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Decrim</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$P_{MAR}$</td>
<td>-5.323*** (0.829)</td>
<td>-0.407*** (0.060)</td>
<td>-5.516*** (0.804)</td>
<td>-0.352*** (0.060)</td>
<td>-0.305*** (0.057)</td>
<td>-0.315*** (0.057)</td>
<td>0.185† (0.130)</td>
<td>0.173*** (0.033)</td>
<td>-0.338*** (0.065)</td>
</tr>
<tr>
<td><strong>Income</strong></td>
<td>-0.018 (0.052)</td>
<td>-0.023 (0.024)</td>
<td>-0.016 (0.050)</td>
<td>-0.013 (0.024)</td>
<td>-0.016 (0.049)</td>
<td>-0.007 (0.024)</td>
<td>-0.009 (0.022)</td>
<td>-0.006 (0.021)</td>
<td>-0.071*** (0.020)</td>
</tr>
<tr>
<td>Age1419</td>
<td>1.963*** (0.418)</td>
<td>1.928*** (0.237)</td>
<td>1.963*** (0.416)</td>
<td>1.960*** (0.240)</td>
<td>1.962*** (0.418)</td>
<td>1.961*** (0.240)</td>
<td>1.952*** (0.205)</td>
<td>1.952*** (0.088)</td>
<td>-0.102</td>
</tr>
<tr>
<td><strong>Age2024</strong></td>
<td>2.084*** (0.409)</td>
<td>2.055*** (0.235)</td>
<td>2.074*** (0.407)</td>
<td>2.101*** (0.237)</td>
<td>2.084*** (0.409)</td>
<td>2.109*** (0.237)</td>
<td>2.105*** (0.202)</td>
<td>2.105*** (0.076)</td>
<td>-0.003</td>
</tr>
<tr>
<td>Age2529</td>
<td>2.022*** (0.404)</td>
<td>2.074*** (0.233)</td>
<td>2.003*** (0.403)</td>
<td>2.124*** (0.235)</td>
<td>2.023*** (0.404)</td>
<td>2.133*** (0.235)</td>
<td>2.113*** (0.076)</td>
<td>2.113*** (0.030)</td>
<td></td>
</tr>
<tr>
<td>Age3034</td>
<td>1.854*** (0.404)</td>
<td>1.874*** (0.231)</td>
<td>1.837*** (0.402)</td>
<td>1.914*** (0.234)</td>
<td>1.854*** (0.404)</td>
<td>1.920*** (0.234)</td>
<td>1.908*** (0.069)</td>
<td>1.908*** (0.007)</td>
<td></td>
</tr>
<tr>
<td>Age3539</td>
<td>1.894*** (0.402)</td>
<td>1.799*** (0.231)</td>
<td>1.876*** (0.401)</td>
<td>1.816*** (0.234)</td>
<td>1.894*** (0.402)</td>
<td>1.822*** (0.234)</td>
<td>1.848*** (0.068)</td>
<td>1.848*** (0.038)</td>
<td></td>
</tr>
<tr>
<td>Age4069</td>
<td>1.113*** (0.398)</td>
<td>1.211*** (0.228)</td>
<td>1.103*** (0.401)</td>
<td>1.233*** (0.231)</td>
<td>1.113*** (0.398)</td>
<td>1.237*** (0.232)</td>
<td>1.212*** (0.067)</td>
<td>1.212*** (0.032)</td>
<td></td>
</tr>
<tr>
<td><strong>Male</strong></td>
<td>0.201*** (0.060)</td>
<td>0.237*** (0.032)</td>
<td>0.199*** (0.059)</td>
<td>0.247*** (0.032)</td>
<td>0.201*** (0.060)</td>
<td>0.250*** (0.032)</td>
<td>0.238*** (0.028)</td>
<td>0.238*** (0.022)</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>-0.428*** (0.080)</td>
<td>-0.468*** (0.044)</td>
<td>-0.416*** (0.080)</td>
<td>-0.489*** (0.043)</td>
<td>-0.429*** (0.080)</td>
<td>-0.493*** (0.043)</td>
<td>-0.470*** (0.037)</td>
<td>-0.470*** (0.006)</td>
<td></td>
</tr>
<tr>
<td>Divorce</td>
<td>-0.052 (0.112)</td>
<td>-0.003 (0.059)</td>
<td>-0.067 (0.111)</td>
<td>0.003 (0.060)</td>
<td>-0.051 (0.112)</td>
<td>0.008 (0.061)</td>
<td>-0.001 (0.053)</td>
<td>0.001 (0.052)</td>
<td></td>
</tr>
<tr>
<td>Widow</td>
<td>-0.561* (0.310)</td>
<td>-0.418*** (0.137)</td>
<td>-0.559* (0.307)</td>
<td>-0.433*** (0.139)</td>
<td>-0.561* (0.310)</td>
<td>-0.435*** (0.140)</td>
<td>-0.441*** (0.127)</td>
<td>-0.441*** (0.079)</td>
<td></td>
</tr>
<tr>
<td>Degree</td>
<td>-0.262*** (0.073)</td>
<td>0.043 (0.038)</td>
<td>-0.262*** (0.073)</td>
<td>0.046 (0.038)</td>
<td>-0.262*** (0.073)</td>
<td>0.047 (0.039)</td>
<td>-0.034 (0.034)</td>
<td>-0.034 (0.018)</td>
<td></td>
</tr>
</tbody>
</table>
Table 2: Coefficient Estimates for Marijuana Smoking and Residential Choice Equations (Continue)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Working Stat</td>
<td>0.307† (0.188)</td>
<td>0.265*** (0.080)</td>
<td>0.318* (0.186)</td>
<td>0.256*** (0.081)</td>
<td>0.306† (0.187)</td>
<td>0.249*** (0.081)</td>
<td>0.257*** (0.075)</td>
<td>0.282*** (0.107)</td>
<td>-0.155* (0.083)</td>
</tr>
<tr>
<td>Aboriginal</td>
<td>-0.308† (0.222)</td>
<td>0.377*** (0.123)</td>
<td>-0.356† (0.219)</td>
<td>0.421*** (0.124)</td>
<td>-0.304† (0.218)</td>
<td>0.438*** (0.124)</td>
<td>0.281*** (0.107)</td>
<td>-0.023† (0.016)</td>
<td>0.319*** (0.107)</td>
</tr>
<tr>
<td># Depchild</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.286*** (0.012)</td>
</tr>
<tr>
<td>State unemployment rate</td>
<td>29.474*** (5.021)</td>
<td>-0.021 (0.478)</td>
<td>31.049*** (4.836)</td>
<td>-0.414 (0.471)</td>
<td>29.338*** (4.832)</td>
<td>-0.650 (0.457)</td>
<td>-0.454 (0.455)</td>
<td>-0.007 (0.077)</td>
<td>-0.468 (0.429)</td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td>-0.011 (0.109)</td>
<td></td>
<td>-0.465*** (0.083)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.408 (0.482)</td>
</tr>
<tr>
<td>$\rho_{10} - \rho_{00}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_{10}$</td>
<td></td>
<td>-0.011 (0.109)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho_{00}$</td>
<td></td>
<td></td>
<td>-0.465*** (0.179)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Treatment Effect (ATE)</td>
<td>0.162</td>
<td>0.183</td>
<td>0.137</td>
<td>0.040</td>
<td>0.037</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: (1) Standard errors are in parentheses, (2) *** significant at 1% (two-tailed test), (3) ** significant at 5% (two-tailed test), (4) * significant at 10% (two-tailed test), (5) † significant at 10% (one-tailed test), (6) The percentage of correct predictions for residential choice equation is found to be 78.87%, (7) Notes (1)-(5) also apply to Table 3.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Model (i): Treatment</th>
<th>Model (i): Control</th>
<th>Model (ii): Treatment</th>
<th>Model (ii): Control</th>
<th>Model (iii): Treatment</th>
<th>Model (iii): Control</th>
<th>Model (iv)</th>
<th>Model (v)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decrim</td>
<td>-2.111*** (0.340)</td>
<td>-0.145*** (0.021)</td>
<td>-2.193*** (0.318)</td>
<td>-0.130*** (0.022)</td>
<td>-2.101*** (0.324)</td>
<td>-0.118*** (0.021)</td>
<td>0.071†</td>
<td>0.067***</td>
</tr>
<tr>
<td>P_MAR</td>
<td>-0.007 (0.021)</td>
<td>-0.008 (0.008)</td>
<td>-0.016 (0.020)</td>
<td>-0.005 (0.009)</td>
<td>-0.006 (0.019)</td>
<td>-0.003 (0.009)</td>
<td>-0.127***</td>
<td>-0.127***</td>
</tr>
<tr>
<td>Income</td>
<td>0.438*** (0.169)</td>
<td>0.307*** (0.092)</td>
<td>0.503*** (0.164)</td>
<td>0.335*** (0.094)</td>
<td>0.433*** (0.048)</td>
<td>0.351*** (0.022)</td>
<td>-0.003</td>
<td>-0.003</td>
</tr>
<tr>
<td>Age1419</td>
<td>0.518*** (0.163)</td>
<td>0.627*** (0.089)</td>
<td>0.451*** (0.163)</td>
<td>0.611*** (0.090)</td>
<td>0.523*** (0.044)</td>
<td>0.599*** (0.026)</td>
<td>0.352***</td>
<td>0.352***</td>
</tr>
<tr>
<td>Age2024</td>
<td>0.515*** (0.161)</td>
<td>0.630*** (0.088)</td>
<td>0.448*** (0.161)</td>
<td>0.613*** (0.089)</td>
<td>0.520*** (0.045)</td>
<td>0.601*** (0.025)</td>
<td>0.598***</td>
<td>0.598***</td>
</tr>
<tr>
<td>Age2529</td>
<td>0.502*** (0.161)</td>
<td>0.603*** (0.087)</td>
<td>0.439*** (0.160)</td>
<td>0.590*** (0.088)</td>
<td>0.507*** (0.049)</td>
<td>0.580*** (0.031)</td>
<td>0.578***</td>
<td>0.578***</td>
</tr>
<tr>
<td>Age3034</td>
<td>0.505*** (0.160)</td>
<td>0.588*** (0.086)</td>
<td>0.442*** (0.160)</td>
<td>0.577*** (0.088)</td>
<td>0.510*** (0.048)</td>
<td>0.567*** (0.034)</td>
<td>0.571***</td>
<td>0.571***</td>
</tr>
<tr>
<td>Age3539</td>
<td>0.385** (0.158)</td>
<td>0.452*** (0.083)</td>
<td>0.349** (0.158)</td>
<td>0.453*** (0.086)</td>
<td>0.388*** (0.098)</td>
<td>0.450*** (0.063)</td>
<td>0.443***</td>
<td>0.443***</td>
</tr>
<tr>
<td>Male</td>
<td>0.078*** (0.024)</td>
<td>0.079*** (0.013)</td>
<td>0.079*** (0.023)</td>
<td>0.086*** (0.013)</td>
<td>0.078*** (0.023)</td>
<td>0.089*** (0.027)</td>
<td>0.085***</td>
<td>0.085***</td>
</tr>
<tr>
<td>Married</td>
<td>-0.161*** (0.032)</td>
<td>-0.144*** (0.018)</td>
<td>-0.164*** (0.032)</td>
<td>-0.157*** (0.016)</td>
<td>-0.161*** (0.028)</td>
<td>-0.163*** (0.013)</td>
<td>-0.157***</td>
<td>-0.157***</td>
</tr>
<tr>
<td>Divorce</td>
<td>-0.021 (0.045)</td>
<td>-0.001 (0.021)</td>
<td>-0.027 (0.044)</td>
<td>0.001 (0.022)</td>
<td>-0.020 (0.044)</td>
<td>0.003 (0.023)</td>
<td>-0.001</td>
<td>-0.0004</td>
</tr>
<tr>
<td>Widow</td>
<td>-0.205* (0.123)</td>
<td>-0.131*** (0.049)</td>
<td>-0.216* (0.122)</td>
<td>-0.142*** (0.022)</td>
<td>-0.204* (0.098)</td>
<td>-0.147*** (0.041)</td>
<td>-0.149***</td>
<td>-0.149***</td>
</tr>
<tr>
<td>Degree</td>
<td>-0.101*** (0.029)</td>
<td>0.016 (0.013)</td>
<td>-0.104*** (0.029)</td>
<td>0.017 (0.014)</td>
<td>-0.101*** (0.027)</td>
<td>0.018 (0.015)</td>
<td>-0.013</td>
<td>-0.013</td>
</tr>
<tr>
<td>Working Status</td>
<td>0.122* (0.074)</td>
<td>0.099*** (0.028)</td>
<td>0.123* (0.074)</td>
<td>0.099*** (0.030)</td>
<td>0.121† (0.074)</td>
<td>0.097*** (0.032)</td>
<td>-0.013</td>
<td>-0.013</td>
</tr>
<tr>
<td>Aboriginal</td>
<td>-0.118† (0.088)</td>
<td>0.143*** (0.044)</td>
<td>-0.141† (0.087)</td>
<td>0.164*** (0.046)</td>
<td>-0.116† (0.080)</td>
<td>0.172*** (0.049)</td>
<td>0.109**</td>
<td>0.110***</td>
</tr>
</tbody>
</table>
In this paper, Working Status is assigned a value of 1 if that person is unemployed and a value 0 otherwise. The coefficient of Working Status is found to be positive and significant for binary probit, bivariate probit, the control group of two-part, restricted and unrestricted endogenous probit switching models. The estimated marginal effect on Table 3 suggests that a reference person can experience up to 12% higher chance of becoming marijuana smoker when he is unemployed. On the contrary, we find no evidence that being unemployed leads to higher prevalence of the drug for the treatment group of model (i) and (iii).

Finally, we turn to the ethnic variable. With models (iv) and (v), we find that being an Aboriginal or Torres Strait Islander has positive and significant effect on participation in marijuana use, but not so for the treatment group of models (i), (ii), and (iii). When allowing for joint dependence between marijuana smoking and residential choice, we find that in decriminalized states respondents with this ethnic origin have more or less the same probability of becoming marijuana smokers as individuals from other ethnic backgrounds.

The estimation results for residential choice equation are provided in Table 2. We find that household income, divorced status, working status, ethnic Aboriginal, and state unemployment rate are important factors for predicting residential choice. These variables are either excluded or not significant for the marijuana behavior equations.

5. PROPENSITY SCORE STRATIFICATION MATCHING

When marijuana smoking equations, $y_{it}^*$ and $y_{0it}^*$, are unspecified, the ATE may still be identifiable and estimable under the assumption that conditional on a set of confounding variables, $w_i = \{x_i, z_i\}$, $(y_{it}^*, y_{0it}^*) \perp d_i$ (ignorable treatment selection, see Heckman and Robb (1985), Rosenbaum and Rubin (1983)). In this section, we use Rosenbaum and Rubin (1983)’s propensity score method to correct for selection on observables, where the propensity score is
defined as the conditional probability of being assigned into treatment given the covariates. In our context, this is simply the conditional probability of residing in decriminalized states given observable variables. Let $w_i = \{x_i \cup z_i\}$. We denote the propensity score $\Pr(d_i = 1 \mid w_i)$ by $p(w_i)$.

Under the assumptions $0 < p(w_i) = \Pr(d_i = 1 \mid w_i) < 1$ and $(y_{i1}^*, y_{0i}^*) \perp d_i \mid w_i$, we have

$(y_{i1}^*, y_{0i}^*) \perp d_i \mid p(w_i)$, \hspace{1cm} (5.2)

and

$w_i \perp d_i \mid p(w_i)$. \hspace{1cm} (5.3)

Equation (5.3) establishes that conditioning on the propensity score, the distribution of covariates $w_i$ must be the same across the treatment and the control groups. In other words, given the propensity score, the assignment into treatment is random. We compute ATE under the assumptions (5.1).

Propensity score stratification matching can be implemented by following these steps: (i) estimating the propensity score either parametrically or non-parametrically, (ii) dividing the propensity score into different intervals such that for each interval there is a presence of both treated and untreated units, (iii) within each stratum, calculating means difference of the treatment and the control outcomes, and finally (iv) computing ATET and ATE by simply taking the weighted average of these differences with the weight being the frequency of treated observations or the frequency of both treated and untreated observations in each interval respectively (Becker and Ichino 2002, p.7 or Cameron and Trivedi 2005, pp. 875-876).

We follow Dehejia and Wahba (1999) and Becker and Ichino (2002) in doing the empirical estimation. First, as shown by Horowitz (1993) and Newey, Powell and Walker (1990) that there is not much difference in predicting the outcomes using parametric or semiparametric methods, we estimate the propensity score by running a binary probit estimation given $w_i$. This
step provides us with the estimated propensity score, $p(w_i)$, which we can use to plot histograms for both the treatment and the control groups. We draw histograms in Table 4 by focusing on a range of propensity score between 0.05 and 0.45 because both the treated and the control units are presented in this region. We proceed to calculate ATE, ATET, and their associated standard errors in Table 5. Four different ranges of overlapping region (i.e. 0.05-0.45, 0.075-0.425, 0.05-

**Table 4: Histograms of Estimated Propensity Scores in the Overlapping Region**

**Figure 1: Treatment group**

![Histogram of Treatment Group](image1)

**Figure 2: Control group**

![Histogram of Control Group](image2)
0.4, and 0.1-0.35) and two different ways of partitioning the propensity scores (i.e. length of interval 0.025 and 0.05) are being considered in this study. In addition to our manual partition, we also use STATA’s `pscore` command to divide propensity scores into smaller strata. The results in Table 5 demonstrate that ATE varies between 0.059 and 0.112, while ATET fluctuates from -0.069 to -0.021. Since both ATE and ATET are highly sensitive to the way in which the propensity score is stratified, we check whether condition (5.3) is met by testing whether there is any difference in the first moment between the treatment and the control groups for each interval. Regardless of how we or STATA partition the propensity score, the t-test and the t-squared test always rejects the null hypothesis of means’ equality between the treatment and the control groups. This result suggests that conditioning on the propensity score the distribution of $w_i$ is different between the treated and the control units, which implies that balancing condition is violated. The violation could be because our sample size is not large enough to perform reliable

**Table 5: ATE and ATET using Propensity Score Stratification Method**

<table>
<thead>
<tr>
<th>Range of Estimated Propensity Score</th>
<th>Number of Treatment observations</th>
<th>Number of Control observations</th>
<th>ATE</th>
<th>ATET</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05 – 0.45 Length of interval 0.025</td>
<td>2810</td>
<td>10301</td>
<td>0.092*** (0.026)</td>
<td>-0.032 (0.053)</td>
</tr>
<tr>
<td>0.05 – 0.45 With STATA interval</td>
<td>2810</td>
<td>10301</td>
<td>0.096*** (0.025)</td>
<td>-0.069* (0.038)</td>
</tr>
<tr>
<td>0.075 – 0.425 Length of interval 0.025</td>
<td>2432</td>
<td>9509</td>
<td>0.074*** (0.018)</td>
<td>-0.056** (0.025)</td>
</tr>
<tr>
<td>0.075 – 0.425 With STATA interval</td>
<td>2432</td>
<td>9509</td>
<td>0.086*** (0.020)</td>
<td>-0.056** (0.025)</td>
</tr>
<tr>
<td>0.05 – 0.4 Length of interval 0.05</td>
<td>2116</td>
<td>10295</td>
<td>0.098*** (0.017)</td>
<td>-0.024* (0.014)</td>
</tr>
<tr>
<td>0.05 – 0.4 With STATA interval</td>
<td>2116</td>
<td>10295</td>
<td>0.112*** (0.026)</td>
<td>-0.026** (0.012)</td>
</tr>
<tr>
<td>0.1– 0.35 Length of interval 0.05</td>
<td>1909</td>
<td>8263</td>
<td>0.067*** (0.016)</td>
<td>-0.021† (0.015)</td>
</tr>
<tr>
<td>0.1– 0.35 With STATA interval</td>
<td>1909</td>
<td>8263</td>
<td>0.059*** (0.020)</td>
<td>-0.022† (0.015)</td>
</tr>
</tbody>
</table>

Note: (i) *** significant at 1%, (ii) ** significant at 5%, (iii) * significant at 10%, (iv) † significant at 10% one-tailed test, (v) standard errors are in the parentheses.
nonparametric estimates or because conditional independence assumption does not hold. Thus, our propensity score matching estimates of the ATE and ATET could capture not only the treatment effect but also the impact of differences in observable and unobservable covariates on smoking outcome.

6. SPECIFICATION ANALYSES

Our numerical analyses using parametric and nonparametric methods yield very different inferences about the impact of decriminalization on marijuana smoking prevalence. In this section, we provide specification analyses to investigate which model or method could be more accurate in capturing the essentials of our data. As discussed in the last section, our propensity score matching analysis fails to satisfy the balancing condition implied by the conditional independence assumption. This failure could be due to the violation of conditional independence assumption. On the other hand, the parametric approach can simultaneously take account of both selection on observables and unobservables provided that the assumptions of the model are consistent with the data generating process. If the parametric form is misspecified, inference based on parametric specification can be misleading. Therefore, we shall first conduct a nonparametric kernel consistent test of the null of unrestricted endogenous probit switching model against the nonparametric alternative that neither places the restriction on the functional form nor makes any distributional assumption. If the endogenous probit switching model is not rejected, we further conduct likelihood ratio tests to choose between the unrestricted endogenous switching model and its nested binary probit, sample selection model, two-part model, and restricted endogenous switching model.

6.1 Parametric versus Nonparametric Modeling
Our main hypothesis for a parametric model is the unrestricted endogenous probit switching model that nests the restricted endogenous switching model, two-part model, sample selection model and dummy variable model as special cases. However, it is possible that the data generating process may follow other alternative specifications. To check the adequacy of our unrestricted endogenous switching model in capturing the essential characteristics of the observed data, we test the null of unrestricted endogenous switching model against the departure from the null in any direction.

The basic idea of Bierens (1982), Hong and White (1995), and others on testing a parametric null of \( y_i = m(x_i, \beta) + u_i \) against the departure from the null in any direction is that under the null \( H_0: E(u_i \mid x_i) = 0 \), while under the alternative \( H_1: E(u_i \mid x_i) \neq 0 \). Testing \( E(u_i \mid x_i) = 0 \) is equivalent to testing

\[
E\{u_i, E(u_i \mid x_i) f(x_i)\} = 0.
\]  

(6.1)

Because our \( x \) contains both continuous and discrete variables, following Hsiao, Li and Racine (2007) and Li and Racine (2007) we use

\[
I_n = \frac{1}{n} \sum_{i=1}^{n} u_i \left[ \frac{1}{n} \sum_{j \neq i} \left[ \prod_{s=1}^{q} \frac{1}{h_s} k \left( \frac{x_{is} - x_{js}}{h_s} \right) \right] ^r \prod_{s=1}^{r} I(x_{is}, x_{js}, \lambda_s) \right]
\]

(6.2)

as a sample analogue of (6.1), where

\[
\hat{u}_i = y_i - E(y_i \mid d_i = 1) = y_i - \Phi \left( \alpha_i + \beta_i x_i + \rho_i \frac{\phi(y_i)}{\sigma_1} \frac{\Phi(y_i)}{\sigma_1} \right) ^{1/2} \] if \( d_i = 1 \), and

(6.3)
\begin{equation}
\hat{u}_i = y_i - E(y_i | d_i = 0) = y_i - \Phi \left( \frac{\phi(y_z i) \left( 1 - \Phi(y_z i) \right)}{\left( 1 - \rho_{0v}^2 \right)^{1/2}} \right) \text{ if } d_i = 0
\end{equation}

\( x^c \) and \( x^d \) denote continuous and discrete variables, \( q \) and \( r \) denote the dimension of continuous and discrete regressors, \( k(\cdot) \) denotes the normal kernel function and \( l(\cdot) \) denotes the kernel function of the form

\( l(x^d_{i+}, x^d_{j+}, \lambda_s) = 1 \text{ if } x^d_{i+} = x^d_{j+}, \text{ and } l(x^d_{i+}, x^d_{j+}, \lambda_s) = \lambda_s \text{ otherwise} \)\hspace{1cm} (6.4)

when \( x^d_i \) is an unordered discrete regressor and

\( l(x^d_{i+}, x^d_{j+}, \lambda_s) = 1 \text{ if } x^d_{i+} = x^d_{j+}, \text{ and } l(x^d_{i+}, x^d_{j+}, \lambda_s) = 2^{1/2} |x^d_{i+} - x^d_{j+}| \text{ otherwise} \)\hspace{1cm} (6.5)

when \( x^d_i \) is an ordered discrete regressor, where \( \lambda_s \in [0,1] \). It is shown by Hsiao, Li and Racine (2007) that under the null,

\( J_n = n \left( h_1 \cdots h_q \right)^{1/2} I_n / \sqrt{\Omega} \rightarrow N(0,1) \)\hspace{1cm} (6.6)

where

\( \Omega = \frac{2h_1 \cdots h_q}{n^2} \sum_{i,j} \sum_{i,j} \hat{u}_i \hat{u}_j \left( \prod_{s=1}^q \frac{1}{h_s} k \left( \frac{x^c_{i+} - x^c_{j+}}{h_s} \right) \right) \left[ \prod_{s=1}^q l(x^d_{i+}, x^d_{j+}, \lambda_s) \right] \). \hspace{1cm} (6.7)

However, under \( H_0 \), \( J_n \) converges to the standard normal distribution at the slow rate of \( O_p \left( (h_1 \cdots h_q)^{1/2} \right) \). To overcome this slow convergence problem, we use the following bootstrapping procedure to approximate the finite sample distribution of (6.6)

(i) From maximum likelihood estimates \( \hat{\beta}_1 \), \( \hat{\beta}_0 \), \( \hat{\alpha}_1 \), \( \hat{\alpha}_0 \), \( \hat{\gamma} \), \( \hat{\rho}_{0v} \), \( \hat{\rho}_{00} \) of the endogenous probit switching model (2.1)-(2.6), we compute \( \hat{p}_{11} = \Pr(y_i = 1 | d_i = 1) \), \( \hat{p}_{01} = 1 - \hat{p}_{11} \), \( \hat{p}_{10} = \Pr(y_i = 1 | d_i = 0) \), and \( \hat{p}_{00} = 1 - \hat{p}_{10} \).
(ii) Generate bootstrapping samples \( \{y_i^b, x_i^b\}_{i=1}^n \) where \( y_i^b \) is the dependent variable from bootstrapping, \( x_i \) is the explanatory variables from the original data, \( n \) is the number of observations that is equal to 14008 in our case, and \( y_i^b \) is randomly drawn from the binomial distribution with \( p = p_i^{11} \) if \( d_i = 1 \) and \( p = p_i^{10} \) if \( d_i = 0 \).

(iii) Use the bootstrapping samples \( \{y_i^b, x_i^b\}_{i=1}^n \) from (ii) to estimate the endogenous probit switching model (i.e. model (2.1)-(2.6)) by the maximum likelihood method and obtain a new set of \( \hat{\beta}_b^1, \hat{\beta}_b^0, \hat{\alpha}_b^1, \hat{\alpha}_b^0, \hat{\gamma}_b, \hat{\rho}_b^{1u}, \) and \( \hat{\rho}_b^{0u} \).

(iv) Use the maximum likelihood estimators \( \hat{\beta}_b^1, \hat{\beta}_b^0, \hat{\alpha}_b^1, \hat{\alpha}_b^0, \hat{\gamma}_b, \hat{\rho}_b^{1u}, \) and \( \hat{\rho}_b^{0u} \) from step (iii) and the bootstrapping samples \( \{y_i^b, x_i^b\}_{i=1}^n \) from step (ii) to compute bootstrapping residual \( \hat{u}_i^b \) and compute \( J_n^b \) using the ad hoc plug-in bandwidth \( h_n = (1.06)\sigma_s(n)^{-1/(2p+1)} \) where \( p \) is the order of the kernel function, \( l \) denotes the dimension of continuous regressors, \( \sigma_s \) is the sample standard deviation for variable \( s \). We set \( \lambda_s = 0 \) for discrete regressors.

(v) We repeat steps (ii) through (iv) 399 times and use the sorted test statistics \( \{J_{nj}^b\}_{j=1}^{399} \) to construct a bootstrap empirical distribution, which in turn is used to approximate the null distribution of the test statistic \( J_n \). We conduct a two-sided test by comparing \( J_n \) to the 1%, 5%, and 10% critical values from the bootstrap empirical distribution.

The bootstrap empirical distribution finds that the two-sided critical values at 1%, 5%, and 10% are 2.721, 1.995, and 1.625, respectively. Our \( J_n \) is equal to 1.282. In other words, it does not
appear that our endogenous probit switching model is contradicted by the information contained in the data.

6.2 Likelihood Ratio Tests for Nested Alternatives

Taking the unrestricted endogenous probit switching model (model (2.1)-(2.6)) as the maintained hypothesis and treating other parametric models as nested hypotheses, it follows that

(a) if

\[ H_0^* : \rho_{1w} = \rho_{0v} \neq 0, \]  
(6.8)

the model becomes the restricted endogenous probit switching model, (b) if

\[ H_0^{**} : \rho_{1w} = \rho_{0v} = 0, \]  
(6.9)

the model has the form of a generalized two-part model, (c) if

\[ H_0^{***} : \beta_1 = \beta_0 \text{ and } \rho_{1w} = \rho_{0v}, \]  
(6.10)

the unrestricted endogenous switching model is reduced to the sample selection model (Amemiya 1985), and (d) if

\[ H_0^{****} : \beta_1 = \beta_0 \text{ and } \rho_{1w} = \rho_{0v} = 0, \]  
(6.11)

the unrestricted endogenous switching model is reduced to the conventional dummy variable approach (model (2.7)). To conduct specification tests for \( H_0^* \), \( H_0^{**} \), \( H_0^{***} \) and \( H_0^{****} \), we compute likelihood ratio statistics \( LR_1 \), \( LR_2 \), \( LR_3 \) and \( LR_4 \), respectively. Each of them is associated with different degree of freedom.

Table 6 provides the results of this specification analysis. The \( LR_2 \) and \( LR_4 \) firmly reject the independence assumption between the errors of marijuana smoking equation and the errors of residential choice equation, while \( LR_1 \) demonstrates that the correlation of these error terms are different for the treatment and the control groups. The \( LR_3 \) and \( LR_4 \) also show that individuals do behave differently if marijuana smoking is decriminalized. In other words, our specification analysis appears to favor the unrestricted endogenous probit switching model.
Table 6: Likelihood Ratio Tests

<table>
<thead>
<tr>
<th>Model under Null Hypothesis</th>
<th>Restricted Switching</th>
<th>Two-Part</th>
<th>Bivariate Probit</th>
<th>Binary Probit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model under Alternative Hypothesis</td>
<td>Unrestricted Switching</td>
<td>Unrestricted Switching</td>
<td>Unrestricted Switching</td>
<td>Unrestricted Switching</td>
</tr>
<tr>
<td>Degree of Freedom of the Chi-Square Dist.</td>
<td>1</td>
<td>2</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>Likelihood Ratio Test Statistics</td>
<td>LR₁ = 5.3**</td>
<td>LR₂ = 8.49**</td>
<td>LR₃ = 80.964***</td>
<td>LR₄ = 80.973***</td>
</tr>
</tbody>
</table>

Note: (1) *** significant at 1% (two-tailed test), (2) ** significant at 5% (two-tailed test)

7. CONCLUSION

This paper uses the 2001 wave of National Drug Strategy Household Survey (NDSHS) to empirically examine the impact of marijuana decriminalization on marijuana smoking prevalence in Australia. Both parametric and nonparametric approaches are adopted. The advantage of nonparametric approach is that no functional form or distributional assumption needs to be imposed. The disadvantage is that the conditional independence assumption (ignorable treatment assignment assumption) is a maintained hypothesis. Furthermore, the impacts of other socio-demographic effects on marijuana smoking are not estimated. The advantage of parametric specification is that both the issues of selection on observables and unobservables can be taken account of and the impact of each variable on the outcomes can be assessed provided the parametric assumptions are not contradicted by the information of the data. The disadvantage is that both functional form and distributional assumptions are imposed. If these assumptions are incorrect, the resulting inferences will be misleading. However, our data analysis appears to indicate that the validity of the conditional independence assumption for nonparametric matching adjustment is contradicted by the information of the data. On the other hand, our parametric model does not appear to be contradicted by the information of the data. As a matter of fact, the world is not that benevolent. We could be asking too much of our data. “We want them to test
our theories, provide us with estimates of important parameters, and disclose to us the exact form of the interrelationships between the various variables” (Griliches 1967, pp. 17-18). When the information contained in the data is limited, an integration of behavioral assumption and parametric specification may allow us to extract more useful information from the data.

Conditional on the state decriminalization decision, our specification analysis appears to favor the unrestricted endogenous probit switching model that takes into account both selection on observables and unobservables. This model suggests that decriminalization policy leads to higher marijuana smoking participation. It indicates that on average living in decriminalized states significantly increases the probability of smoking marijuana by 16.2%. This estimate is larger than those obtained from the dummy variable approach, the sample selection model, the two-part model, and other existing studies. The discrepancies could be due to the use of different data sets. It could also be due to different model specifications as demonstrated in this paper. However, our specification analysis appears to favor the model that allows both the endogeneity of the decriminalization dummy and the flexibility of different behavioral patterns due to changes in legal or institutional environment.

Acknowledgements

The authors would like to thank Ragui Assaad, Deborah Levison, and Cristobal Ridao-Cano for making available their maximum likelihood computer program. We also thank two referees, an editor, and John Ham, Qi Li, Chiahui Lu, Jeffrey Nugent, Michael Nichols, John Strauss, Vai-lam Mui for their helpful comments. We are also grateful for Jeffrey Racine for making available his R-program to conduct the specification analysis.

References


