Illegal Drugs, Education, and Labor Market Outcomes*

Alvaro Mezza (Federal Reserve Board) Moshe Buchinsky (UCLA)

January, 2014

Abstract

In this paper we investigate the causal effects of consuming illegal drugs on educational attainment, employment and wages. To identify these effects, we develop and estimate the first dynamic structural model to jointly consider decisions of whether to consume drugs, attend school, participate in the labor force, and save.

Using data from the National Longitudinal Survey of Youth 1997 (NLSY97), we focus our analysis on males; the period of analysis begins at age 13, when they are young enough to have had no experience with drugs. Contrary to findings in the literature, non-drug users have higher wages than marijuana and/or hard drug users. This effect is small for individuals who consume marijuana in low doses but increases with the frequency of drug use. Results from a counterfactual experiment suggest that a 30 percent increase in the price of marijuana each period would reduce the number of marijuana consumers among the 13- to 29-year-olds by 23 percent. Individuals who are dissuaded from consuming marijuana due to the higher price would also increase their annual income by an amount, on average, of $2,763 by age 29. The reduction in drug use also has a positive effect on school attendance and employment for this group.

*E-mail: alvaro.a.mezza@frb.gov. We want to thank Maria Casanova, Doug McKee, Jinyong Hahn, Maurizio Mazzocco, Guillermo Ordoñez, Claudia Ruiz-Ortega, Emily Wiemers and Hernan Winkler for their valuable suggestions. The analysis and conclusions set forth are those of the authors and do not indicate concurrency by other members of the Federal Reserve System. All remaining errors are of our own.

1 Introduction

Although the US government has long been concerned about the effect of drug consumption on workplace productivity, the recent spike in drug use among youth raises new and urgent concerns about its effect on education and labor market outcomes.[1] The US office of National Drug Control Policy (ONDCP) reported that the 30-day prevalence of drug use among youth aged 12 or older increased 8.75 percentage points during the period 2007-2009. The spike in drug consumption among the youth suggests that young people are increasingly making decisions about schooling and work simultaneously with the decision to consume drugs. Since drug users are observed to acquire less education than non drug users, the recent trend calls for a fresh look at the linkages between drugs consumption, educational attainment and labor market outcomes.[2] In light of the Obama administration's 2010 recent policies to significantly reduce drug consumption among youth, it is especially important to shed some light on the effects of drug use on education and labor market outcomes.[3]

By developing and estimating the first structural dynamic model incorporating simultaneous decisions on drug consumption, school attendance, and participation in the labor market, and allowing these decisions to affect future opportunities, this paper makes a first attempt at capturing the simultaneity and dynamics that characterize these decisions. Incorporating these factors in a single framework allows us to directly examine and better understand the causal effects of drug use. It allows us to consider counterfactual experiments that would be impossible to conduct under the so-called reduced-form approach, such as: (i) examine the effect of changes in drug prices on education, labor market outcomes, and consumption of soft and hard drugs; (ii) examine the potential effects of educational costs on drug use and the consequent choices of education and labor market outcomes, and (iii) examine the effect of drug use while in adolescence in labor market outcomes and on the educational achievements of these young individuals.

There are several obstacles to establishing a causal link between drug use and education or drug use and labor market outcomes. First, while substance abuse can lead to employment problems, employment problems can also induce drug use. Second, since drugs are considered a normal good, having a higher income may increase drug consumption with no apparent effect on performance in the short run. Third, unobserved factors, such as levels of ability, which affect drug consumption and employment simultaneously, could create a spurious relation between employment and drug consumption. Finally, according to the theory of human capital accumulation, investments in education and job experience are dynamic in

[1] Concerns about workplace productivity, absenteeism, and safety led the government to enact the federal law known as Drug-free workplace act of 1988. This law requires some federal contractors and all federal grantees to maintain a workplace environment free of illegal drugs. Though not stipulated by the law many federal contractors started with employee testing programs for drug consumption and many privately owned companies followed the initiative. Shepard and Clifton (1998) estimate that in the mid-1990s the U.S. spent over $1 billion annually on testing about 20 million workers for drug use.

[2] According to data from the National Longitudinal Survey of Youth 1997 (NLSY97), which is a nationally representative sample of the U.S. population born during the years 1980 through 1984, drug users have almost one year less of completed education by age 24 than non drug users.

[3] Some of Obama's administration goals reported in the 2010 National Drug Control Strategy are: (i) reduce 30-day prevalence of drug use among 12 to 17 year olds by 15% and (ii) reduce the 30-day prevalence of drug use among young adults aged 18 to 25 by 10% by the year 2015.

nature. Dynamics are also a key factor in the various papers that have dealt with addiction theories: that is, drug consumers' past consumption determines current and future behavior (for example, Becker and Murphy 1988, Orphanides and Zervos 1995,1998, Gruber and Koszegi 2001, Bernheim and Rangel 2004). Failure to account for any of these factors could bias results in ways that cannot be predicted ex-ante.

Findings in the literature suggest that using drugs has a small negative effect on employment, or no effect at all, and a positive, or zero effect on wages. In terms of education, the evidence is divided among studies that support a causal effect of drugs on educational achievement (see, for example, Chatterji 2006 and Ellickson et al. 1998), that support the reverse effect of educational achievement on drug use (see, for example, Fergusson and Horwood 1997), and that attribute observed effects to unobserved factors affecting drug consumption and education simultaneously (McCaffrey et al. 2008 and Barnes et al. 2005). A possible explanation for the mixed findings is that the current literature, which is based largely on reduced form approaches, is not able to control for the simultaneity of decisions of drug consumption, education, and employment. Studies that consider the effect of drugs on employment and wages usually assume that education and work experience are exogenous. This assumption is problematic, especially when studies consider the behavior of adolescents and young adults since the decisions of whether to work and study are taken simultaneously. Another factor is that the majority of these analyses are based on cross-sectional data, which does not allow them to control for unobserved factors. Finally, these types of studies do not account for the forward looking nature of drug use, limiting their ability to address welfare implications. To address the problems in the existing literature we extend the addiction models of Becker and Murphy (1988) and Bernheim and Rangel (2004) by explicitly incorporating individuals' education and employment decisions within a dynamic finite horizon model of drug consumption choice. In this way, we recognize the simultaneous and dynamic nature of human capital, of work and of drug consumption decisions, endogenizing the accumulation of education, of work experience, and of drugs in individuals' system.

Becker and Murphy's model of addiction is the standard theoretical model of addiction in economics, where forward-looking and time consistent individuals decide whether or not to consume drugs to maximize their expected discounted utility. Various studies have tested this theory empirically, finding support for the assumption that people are forward-looking.[4] However, some concerns have been raised about the validity of the joint assumptions of rationality and time consistency (see for example, Gruber and Koszegi 2001 and Bernheim and Rangel 2004). To account for the possibility that addicts do not behave as assumed by Becker and Murphy, we also extend Bernheim and Rangel's model of addiction, a model that finds support in evidence from neuroscience, psychology, and clinical practice. In Bernheim and Rangel's model individuals are forward-looking but substance use sensitizes addicts to environmental cues that trigger mistaken usage. A specific case of Bernheim and Rangel's model is Becker and Murphy's model in which addicts do not commit mistakes.

The main findings of this paper are as follows. Preliminary estimates obtained by extending Becker and Murphy's model of addictions suggest that the main effects of drugs come

[4] See Grossman et al. (1998) for a summary of empirical applications that test the rationality assumption, and Gordon and Sun (2009) and Choo (2000) for empirical applications that have assumed rational and time consistent behavior in a dynamic framework to study cigarette consumption.

through wages. Contrary to findings in the literature, non-drug users have higher wages than marijuana and/or hard drug users.[5] This effect is small for individuals who consume marijuana in very low doses and increases with the frequency of marijuana and/or hard drug consumption. Results from a counterfactual experiment suggest that a 30 percent increase in the price of marijuana each period would reduce the number of marijuana consumers among 13-29 year-old by 23 percent on average. Individuals who are dissuaded from consuming marijuana due to the higher price would make, on average, $2,763 more annually by age 29. School attendance and employment are also positively affected for this group, though the effects are milder.

The rest of the paper is organized as follows. The next section reviews the relevant literature and puts the current paper into context. Section 3 presents the model and outlines the estimation procedure. Section 4 describes the data. Section 5 discusses results from the estimation the fit of the model and presents counterfactual experiments. Section 6 summarizes and concludes.

2 Background and Relevant Literature

This paper combines two strands of literature: a literature that studies the effect of drugs on labor market outcomes and education, and a literature that studies the role of addictions.

2.1 Labor Market Outcomes, Education and Illegal Drugs

Research on the effect of drugs on the labor market has centered on the effect of drugs on labor force participation, employment and wages. The stylized fact that drug users tend to acquire less education than their counterparts who do not use drugs has also attracted attention (see Tables B.4 and B.5 in appendix B). Despite research efforts, the evidence of the effect of drug use on labor market outcomes and education is inconclusive.

Even though there is a preconception that the detrimental physical and psychological effects of drug use would lead to lower wages and lower employment rates, the literature suggests a positive association or, at worst, no association between drugs and wages; it shows no association, or a small association, between drugs and employment.[6]

In cross-section analyses, Kaestner (1991) and Gill and Michaels (1992) find a positive association between drug use and wages.[7] Register and Williams (1992) reaches the same conclusion when analyzing wages for marijuana users, but finds no association between wages

[5] Hard drugs include cocaine, heroin, LSD, and other drugs different from marijuana that are not prescribed by a physician.

[6] All the results reported in this section refer to the effect of drugs for males in the U.S. However, many authors have analyze the association of drugs with employment and wages for other countries, reaching similar conclusions (see for example, Van Ours (2006) for a study about Holland, and MacDonald and Pudney (2000) for a study about the U.K.). The same holds for studies that have included women in their analyses.

[7] The studies and results mentioned in this section are representative of the results found in a large body of literature on the effect of drugs on wages, employment and education. For a comprehensive analysis of the literature on the association of drugs with employment and wages, see DeSimone (2002) and Van Ours (2006); for drugs and education, see McCaffrey et al. (2008).

and cocaine use. Buchmueller and Zuvekas (1998) also finds a positive association between wages and marijuana drug use for young workers. However, the authors find a negative relationship between wages and drug use for heavy users. Kaestner (1994a) extends his previous cross-sectional results with a panel data analysis. He concludes that once fixed effects are included to control for unobserved heterogeneity, the positive association between wages and drug use disappears. Conti (2009) jointly models cognitive ability, marijuana use and wages using an econometric factor model (see Heckman et al. 2006). The author concludes that once one controls for cognitive ability the positive association between wages and marijuana use disappears.

In cross-section analyses on drug use and employment, Kaestner (1994b), Gill and Michaels (1992), and DeSimone (2002) find a negative association between these variables. Register and Williams (1992) reach the same conclusion when analyzing employment for marijuana users, but the authors find no association between employment and cocaine use. Buchmueller and Zuvekas (1998) also finds a negative association between employment and problematic drug use for people age 30-45 year-olds, but no association for younger people or non problematic users.[8] French et al. (2001) reach a similar conclusion, finding that employment is negatively related to chronic drug use, but unrelated to non chronic use.[9] Kaestner (1994b) extends cross-sectional results using panel data, and concludes that the negative relationship disappears once one controls for fixed effects.

Studies by Ringel et al. (2006) and Burgess and Propper (1998) examine the effect of early drug use on labor market outcomes later in life. They reason that if illegal drug use affects motivation, cognitive skills and/or attention negatively, it would be more harmful during adolescence when decisions about schooling and entering in the labor market are made. Ringel et al. conclude that there is a negative relationship between the consumption of marijuana and later earnings. However, the analysis is based on OLS results so the finding is not necessarily causal. Burgess and Propper find that consumption of soft drugs at younger ages does not affect labor market participation ten years later.

One common problem faced by the majority of these analyses is controlling for the simultaneity in the determination of wages, employment and drug consumption decisions. As DeSimone (2002) and Van Ours (2006) note, a possible explanation for the mixed results found in the literature could be the difficulty in finding reliable instruments to control for the simultaneity.[10]

In the case of research that investigates the effect of illegal drug consumption on education, results are mixed. Many papers have demonstrated the existence of a positive association between illegal drug use and poor educational outcomes. Some papers have interpreted this association as evidence of a causal effect of drugs on educational achievement (see, for example, Chatterji 2006 and Ellickson et al. 1998). However, as McCaffrey et al. (2008) note, these studies have not satisfactorily dealt with the issue of endogeneity. The instruments used in these studies are so weak that they usually fall back on traditional OLS methods for drawing conclusions, which limits interpretation of the results to correlations

[8] Problematic use is indicated by a diagnosis of pathological use, dependence, or daily use.

[9] The term "chronic drug user" refers to those individuals who declare that they consume illegal drugs weekly or more often

[10] DeSimone (2002) and Van Ours (2006) present a detailed discussion about instruments used in the literature, as well as about papers that relied on OLS results.

rather than causal effects. Other papers support the reverse effect, namely that of educational achievement on drug use (see, for example, Fergusson and Horwood 1997). These studies find evidence that drug use precedes poor educational attainment. However, finding evidence that poor educational attainment precedes drug use does not rule out the possibility that drug use negatively affects education. Finally, other papers find evidence of unobserved factors that affect illegal drug consumption and education simultaneously (McCaffrey et al. 2008 and Barnes et al. 2005).

To overcome problems in the existing literature and account for the dynamics of simultaneous decisions on education, employment and drug consumption, we develop and estimate the first structural model incorporating individuals' education and employment decisions within a dynamic finite horizon model of drug consumption choice. By explicitly addressing the simultaneity and the dynamics that characterize these decisions, we are able to study the nature of the selection process over individuals' life-cycles and to better understand the causal effect of using drugs on the variables of interest.

2.2 Rationality and Addictions

Before Becker and Murphy (1988) introduced the concept of rationality in analyzing addictive behaviors, the literature assumed individuals were not rational. Myopia, shifting preferences, and lack of self control were assumed to explain the impulsive behavior associated with addictions (see, for example, Winston 1980, Thaler and Shefrin 1981, Akerlof 1991).

Becker and Murphy (1988), build on Stigler and Becker's (1977) model of addiction by assuming that forward-looking individuals with stable preferences decide whether to consume a potentially addictive good in order to maximize their expected discounted utility, anticipating the future consequences of their choices. Under these assumptions, they explain typical behaviors observed in addicts, such as cycles over time in the consumption of a good (i.e., binges) and giving up a habit or addiction all at once (i.e., cold turkey) as a way to end a strong addiction. In the model, a good is addictive if its past consumption affects its current consumption positively. These complementarities between past and current consumption cause some steady states to be unstable, where small deviations in consumption can lead to large cumulative rises over time in addictive consumption or to rapid falls and abstention. The theory also predicts that consumers adjust their consumption of addictive goods to future changes in prices and that they respond more to permanent changes in prices of addictive goods than to temporary ones.

Many studies have tested this theory empirically by analyzing different behaviors that can become addictive, like consumption of alcohol, cigarettes, coffee or cocaine. These studies analyze first order conditions that prices and quantities need to satisfy, given individuals' utility functions and find that consumption of addictive goods responds to lagged, current, and future price changes as predicted by the rational addiction theory. Arcidiacono and Sieg (2007) use a different approach to test whether rational addiction is a good representation of the data. They apply a dynamic programming approach for smokers and heavy drinkers over 50 years old. They study whether individuals rationally update their consumption behavior as they experience negative health shocks which occur, at least partially, due to smoking and heavy drinking in the past. If people are rational, one expects to observe that individuals reduce their consumption of addictive goods after suffering a negative health shock associated

with consumption of that good. They conclude that forward looking models fit the data better than myopic models.

A limitation of these empirical analyses is that they are not able to distinguish between rational and time consistent consumers, as in Becker and Murphy's theory, from rational but time inconsistent consumers or rational consumers who can commit mistakes.

Gruber and Koszegi (2001) were the first to point this out. Following Laibson's (1997) theory of quasi-hyperbolic discounting, they introduced the idea of time inconsistency in a model of addiction. This introduction intends to capture the idea that agents might have self-control problems regarding the consumption of addictive goods. In this approach, individuals are rational, but their preferences are time inconsistent. This inconsistency arises from allowing the discount factor applied between consecutive future periods to be larger than the one applied between the current period and the next one. Because people are still rational their qualitative reaction to future prices is identical to that predicted by Becker and Murphy. Therefore, the existing empirical literature offers a test of rationality vs. irrationality but not of time consistency vs. time inconsistency.

Other authors, such as Gul and Pesendorfer (2001), Laibson (2001), Bernheim and Rangel (2004) have made important contributions to the theory of addictions with models that depart from Becker and Murphy's approach.

Bernheim and Rangel's (2004) model intends to capture that use among addicts is frequently a mistake. In their model, individuals are rational, but experience with an addictive substance sensitizes them to environmental cues that trigger mistaken usage. The authors argue that these assumptions about addictive behavior find strong support in evidence from neuroscience, psychology, and clinical practice. According to researchers in these sciences, an important aspect of behavior among addicts is that they often see past use as a mistake. According to Bernheim and Rangel, addicts report that they would have been better off in the past as well as in the present if they had not had consumed. They characterize succumbing to cravings as mistakes. Bernheim and Rangel note that none of the other existing theories of addiction are able to explain this important aspect of behavior among addicts.

Bernheim and Rangel also note that prior to the 1990s, neurological theories of addiction were based on the idea that people begin consuming drugs to enjoy the high they provide, and continue consuming drugs to avoid the costs associated with withdrawals and cravings. This view of addictive behavior is at the core of Becker and Murphy's model. However, new evidence from neuroscience suggests that addictive goods interfere with the neural system, causing a specific learning process to malfunction. This learning process, called the "hedonic forecasting mechanism" by Bernheim and Rangel, learns through experience. With experience, the hedonic forecasting mechanism learns to forecast the expected pleasure or pain associated with situations and actions. Addictive substances interfere with the proper operation of this mechanism, causing it to grossly overstate the expected pleasure from consuming the addictive substance, creating a powerful impulse to consume it. To incorporate this into a tractable mathematical model, Bernheim and Rangel introduce the possibility of stochastic state-dependent mistakes in an otherwise dynamic programming problem model of rational addiction.

As in Becker and Murphy's model, the individual is initially drawn to an addictive substance because it provides utility through the experience of a "high". However, once one has experienced the addictive substance, each period one can operate in two modes: a cold mode

in which one rationally chooses the best decision among all alternatives, and a hot mode in which one invariably consumes the addictive substance, irrespective of one's underlying preferences. The likelihood of entering a hot mode is assumed to increase with the level of past consumption, reflecting the higher probability of encountering cues that trigger it.

An interesting feature of Bernheim and Rangel's model is that Becker and Murphy's model is a special case of it, where individuals are not allow to commit mistakes. This feature will be used in this paper to test one theory against the other.[11]

3 The Model

3.1 Baseline model: Adaptation of Becker and Murphy's Rational Addiction Model

In this subsection we adapt Becker and Murphy's model of rational addiction to consider the decisions of whether to consume soft and/or hard illegal drugs. We also extend their model to incorporate the decisions of whether to study, to work and to save to analyze the effect of drugs on educational attainment and labor market outcomes.

3.1.1 The individual's objective

Each period rational forward looking individuals make decisions about whether to work, consume soft and/or hard drugs, invest in human capital, and save, to maximize the discounted present value of their utility:

$$\sum_{t=t_0}^{T} \delta^{t-t_0} E[U^i(c_{it}, l_{it}, a_i^{sdt}, a_i^{hdt})|z_{it}], \tag{1}$$

subject to the following constraints:

$$\begin{aligned} l_{it} &= \overline{L} - h_{it}^w d_{it}^w - h_{it}^s d_{it}^s, \\ B_{it} &= R_t[B_{it-1} + NI_{it} + w_{it} h_{it}^w d_{it}^w - c_{it} - p_t^{sd} a_{it}^{sd} - p_t^{hd} a_{it}^{hd} - cs_{it} d_{it}^s], \\ B_{it} &\geq \underline{B}, \\ S_{it} &\leq 20, \end{aligned} \tag{2}$$

where z_{it} denotes the state vector defined below. $U^i(.)$ denotes the single-period utility for individual i, which is a function of a composite consumption good c_{it} with price normalized to one, leisure, l_{it}, and stock of consumption of soft and hard drugs, a_i^{sdt} and a_i^{hdt}, respectively.

The indicator variable d_{it}^w (or alternatively d_{it}^s) takes value of one if the person works (studies) in period t, and 0 otherwise. The total amount of leisure is given by the amount of waking hours in a year $\overline{L}$ (given by 52x7x16=5824, assuming 8 hours of sleep per day) minus the hours dedicated to work, h_{it}^w, and study, h_{it}^s.

[11] Many recent papers have followed a similar approach to test models of rational behavior with stable preferences against models of rational behavior with unstable preferences (hyperbolic discount model) using dynamic structural models. See for example, Laibson et al. (2007), Fang and Silverman (2009), Fang and Wang (2009), Machado and Sinha (2007), and Shui and Ausubel (2005).

The variable w_{it} denotes the hourly wage when one is employed; cs_i represents the monetary costs of schooling; R_t is the gross return on net wealth, B_{it}, carried over to the next period, and NI_{it} is non-earned income. We assume that the level of borrowing cannot go below a certain minimum level of $\underline{B}$. The budget constraint is affected by expenditure on soft and hard drugs, $p_t^{sd}a_{it}^{sd}+p_t^{hd}a_{it}^{hd}$, where p_t^{sd} (or alternatively p_t^{hd}) is the price of soft (hard) drugs and a_{it}^{sd} (or alternatively a_{it}^{hd}) is the amount of soft (hard) drugs consumed in period t.

The number of years of education, S_{it} has an upper bound of 20, which is the maximum amount of education a person is allowed to report in the database used in this analysis (National Longitudinal Survey of Youth 1997).

The functional form for the single-period utility function is:[12]

$$U^i(c_{it}, l_{it}, a_i^{sdt}, a_i^{hdt})=$$

$$\frac{c_{it}^{\alpha_c}}{\alpha_c}+\frac{l_{it}^{\alpha_l}}{\alpha_l}+\{-\alpha_{sd}\exp[(\sum_{p=t-1}^{T_0}(\beta_{sd})^p a_{ip}^{sd}-a_{it}^{sd})/\alpha_{sd}]+\varepsilon_{it}^{sd}I(a_{it}^{sd}>0)\}$$

$$+\{-\alpha_{hd}\exp[(\sum_{p=t-1}^{T_0}(\beta_{hd})^p a_{ip}^{hd}-a_{it}^{hd})/\alpha_{hd}]+\varepsilon_{it}^{hd}I(a_{it}^{hd}>0)\}$$

where $\alpha_v > 0$, $I(.)$ is the indicator function, and $0 \leq \beta_v < 1$, $c_t > 0, l_t > 0$, $a_{it}^v \geq 0$, with $v = sd, hd$.

The first two terms represent utility from consuming a composite good and leisure. Utilities from consumption and leisure display constant relative risk aversion with relative risk aversion coefficients α_c and α_l. The third and fourth terms represent direct utility from soft and hard drug consumption. We use a constant absolute risk aversion utility function for the consumption of soft and hard drugs. Constant relative risk aversion functions have the property that the marginal utility of consumption is infinite when the amount consumed is zero, implying that an individual will consume positive amounts of the good if he/she has the income to buy it. To avoid the "desirability" that this type of function would impose, we choose constant absolute risk aversion functions for consumption of soft and hard drugs. In a constant absolute risk aversion function the marginal utility of consumption is finite for a zero amount, implying that zero consumption of drugs can be optimal even when income is positive.

Addictions imply linkages in consumption of the same good over time. To differentiate potentially addictive goods like drugs from non-addictive goods, the literature that assumes rationality (for example, Becker and Murphy 1988, Gruber and Koszegi 2001) and the habit persistence literature (Pollak 1970) introduces the stock of past consumption in the utility function. Similarly, we also account for past consumption. T_0 represents the age when individual i starts consuming drugs and the utility from drugs is affected by the whole

[12] In this specification, the utility is separable in c_t, l_t, a_{it}^s, and a_{it}^h. We do not allow for the possibility that drug users derive different utility from c_t and l_t than non-drug users. This restriction will be relaxed in a future version of this paper.

history of past consumption. β is restricted to take values in the interval [0,1) since the effect of past consumption diminishes over time.[13]

By including past consumption of drugs the model captures three characteristics of addictive consumption: withdrawal, tolerance and reinforcement. Withdrawal is the negative physical reaction associated with the cessation or interruption of consumption. Withdrawal is observed if a reduction in drug consumption reduces total utility: $U_{a_t} > 0$ (where U_x represents the first derivative of U with respect to x). Tolerance implies that a higher stock of past consumption yields less satisfaction for a given level of current consumption. That is, tolerance occurs if utility is negatively affected by greater cumulative past consumption: $U_{a_{t-1}} < 0$. Finally, reinforcement means that greater past consumption of a good raises the marginal utility of current consumption, which occurs if $U_{a_t a_{t-1}} > 0$ (i.e., if $\alpha_v > 0$, with v=s,h in the model). This last characteristic is a necessary condition provided by Becker and Murphy for a rational forward-looking consumer to develop a habit.[14]

Finally, the level of utility is also affected by an i.i.d. shock, ε_{it}^v (with v=sd,hd), if current consumption of the drug is positive. Thus, past drug consumption has a known, deterministic effect on the body. Current consumption, on the other hand, is also affected by a stochastic component, reflecting that each experience with drugs is different. This difference can be due to difference in the potency of drugs or due to external conditions (for example, being at a party with friends, peer pressure, etc.) that can enhance or reduce the pleasurable effects of the drugs. This stochastic shock can start the individual on a path of drug consumption.

3.1.2 The Choices

There are five choices that individuals make in each time period:

- whether to attend school
- whether to work
- whether to consume soft and/or hard drugs
- how much to save.

None of these choices is exclusive. A person can work while studying, and can consume soft drugs and hard drugs at the same time.

Schooling Choices

In each period individuals choose whether or not to attend school. Since the decision process starts at age 13, the first educational decision individuals face is to about whether

[13] In the terminology of the habit persistence literature, past consumption enters the utility function when there is a necessary quantity of the good to be consumed. This necessity could be due to physiological and/or psychological reasons.

[14] Bernheim and Rangel's theory of addiction differs on this point from other theories based on rationality. In Bernheim and Rangel's model, reinforcement is not a necessary condition. The possibility that addicts can commit mistakes once they have experience with a substance is enough for a substance to become addictive, even in the absence of reinforcement.

to continue attending primary school or high school, depending on the level of education reached at age 13.

Individuals attending school face some probability of not advancing a grade. This probability depends on work status, whether the person attended school last period or not, highest grade attained, age, current and past drug consumption, and individual characteristics such as race, mother's level of formal education, living in a single-parent household at the beginning of the decision period, and schooling at age 13. Specifically, we parameterize this probability with the following logit function:

$$\Pr(\text{repeating a grade}|z_{it}, d^w_{it}, a^{sd}_{it}, a^{hd}_{it})$$

$$= \frac{\exp(\eta_o+\eta_1 d^w_{it}+\eta_2 d^s_{t-1}+\eta_3 S_{it}+\eta_4 Age_{it}+\eta_5 a^{sd}_{it}+\eta_6 a^{hd}_{it}+\eta_7 \sum_{p=t-1}^{T_0} a^{sd}_{ip}+\eta_8 \sum_{p=t-1}^{T_0} a^{hd}_{ip}+\eta_9 R_i+\eta_{10} MEd_i+\eta_{11} HH+\eta_{12} Ed13)}{1+\exp(\eta_o+\eta_1 d^w_{it}+\eta_2 d^s_{t-1}+\eta_3 S_{it}+\eta_4 Age_{it}+\eta_5 a^{sd}_{it}+\eta_6 a^{hd}_{it}+\eta_7 \sum_{p=t-1}^{T_0} a^{sd}_{ip}+\eta_8 \sum_{p=t-1}^{T_0} a^{hd}_{ip}+\eta_9 R_i+\eta_{10} MEd_i+\eta_{11} HH+\eta_{12} Ed13)}$$

where MEd is mother's education, HH is household type, R is race and Ed13 is education at age 13.

Primary and secondary education are assumed to be free of monetary cost. Post secondary education requires some direct cost, specifically:

$$cs_{it} = cs^u I(12 \leq S_{it} \leq 16) + cs^g I(S_{it} > 16) + \varepsilon^s_{it} I(S_{it} \geq 12),$$

where cs^u (or alternatively cs^g) is the monetary cost of undergraduate (graduate) studies (including tuition cost, books, etc.),and ε^s_{it}, is an i.i.d. shock.

Attending school implies a cost in terms of utility given by the reduction of leisure time. We assume that during the 180 days in an academic year, high school students study for 6 hours a day, and college or graduate students study for 8 hours a day. In this model individuals can stop studying and reenter school at a later period. In reality, it is observed that when people stop attending school, they rarely go back to study later in life. To capture this stylized fact we introduce a cost associated with being out of school for a while. This cost of reentry is faced when the individual did not attend school the previous period. These indirect costs are:

$$h^s_{it} = \gamma_1 + \gamma_2(1 - d^s_{t-1})$$

where γ_1 denotes the number of hours an individual that attended school in the previous period studies, and γ_2 represents the additional hours a person who was not in school the previous period must study.

Working Choice and Reward from Working

At age 16, people start making decisions about work. In this model, a person decides to work full-time, which is 2000 hours a year at 8 hours of work per day over a five-day week, or not work at all. Thus, the number of hours a person decides to work is given by h^w_{it} hours, which is set to 2000.

The reward function for people who work is given by the hourly wage offered:

$$w_{it} = r^* e(z_{it}),$$

where r^* represents the price of a unit of skill in the job and the quantity $e(z_{it})$ represents the individual's total number of productive units. More specifically,

$$\begin{aligned} w_{it} &= \exp\{\beta_1 + \beta_2 a_{it}^{sd} + \beta_3 a_{it}^{hd} + \beta_4 \sum_{p=t-1}^{T_0} a_{ip}^{sd} + \beta_5 \sum_{p=t-1}^{T_0} a_{ip}^{hd} \\ &+(\beta_6 + \beta_7 a_{it}^{sd} + \beta_8 a_{it}^{hd} + \beta_9 \sum_{p=t-1}^{T_0} a_{ip}^{sd} + \beta_{10} \sum_{p=t-1}^{T_0} a_{ip}^{hd}) S_{it} \\ &+(\beta_{11} + \beta_{12} a_{it}^{sd} + \beta_{13} a_{it}^{hd} + \beta_{14} \sum_{p=t-1}^{T_0} a_{ip}^{sd} + \beta_{15} \sum_{p=t-1}^{T_0} a_{ip}^{hd}) X_{it} \\ &+\beta_{16} X_{it}^2 + \varepsilon_{it}^w \end{aligned} \tag{3}$$

where $\beta_1 = \alpha + r$, $r = \log r^*$, S_{it} is years of education, X_{it} is years of experience and ε_{it}^w is an i.i.d. shock.

The function w_{it} takes the form of a Mincer equation that depends on the level of schooling, the level of experience, the history of soft and hard drug consumption and an idiosyncratic shock to the skills of an individual. The returns to education and experience differ between non-users and users of soft and/or hard drugs. Even though two persons can have the same amount of education (experience), the education (experience) they acquire is not identical because of drug use. Investments can be more or less profitable depending on the signs of the β coefficients.

Soft and Hard Drugs Choices

Drug consumption usually begins in early adolescence, so the period of analysis starts at age 13 (see Data section for age of initiation into drug consumption). At this age, individuals start deciding whether or not to consume marijuana and/or other drugs. The amount consumed is determined by the median consumption of that drug observed in the data.

Saving Choices

Individuals are assumed to start deciding how much to save at age 16. Before that, individuals have access to money by transfers (non-earned income). We also include a minimum consumption level, which we estimate, to avoid cases of no consumption when non-earned income is zero.

In order to limit the number of saving choices, we discretize it to four possible categories: the first category denotes a person in debt, the second denotes a person with no savings, and the last two represent two different categories of positive savings.[15]

Initial Conditions and Unobserved Heterogeneity

When individuals are first observed at age 13, they are heterogeneous in several dimensions. To take this into account, we allow individuals to differ in the following observable initial conditions: race, mother's level of formal education, non-earned income at age 13, schooling at age 13, and living in a single-parent household at the beginning of the decision

[15] For more details, see the Data section.

period. However, other sources of heterogeneity that are not captured by these observed initial conditions can be the reason people take different decisions in similar conditions. To control for the existence of unobserved heterogeneity conditional on the observed initial conditions among individuals, the model incorporates a finite number of individual types. This approach of controlling for unobserved heterogeneity in a dynamic discrete choice framework was first introduced by Keane and Wolpin (1997) and is commonly used in the literature.[16]

Distributional Assumptions

The model contains 4 idiosyncratic shocks: ε_{it}^{w} that comes from the wage equation; ε_{it}^{s} that comes from the cost of schooling equation; and ε_{it}^{sd} and ε_{it}^{hd} that come from the utility of soft and hard drugs, respectively. This information is summarized in the following vector:

$$\varepsilon_{it} = (\varepsilon_{it}^{w}, \varepsilon_{it}^{s}, \varepsilon_{it}^{sd}, \varepsilon_{it}^{hd})'.$$

It is assumed that ε_{it} follows a multivariate normal distribution such that:

$$\varepsilon_{it} \sim N(\mu, \Sigma_{\varepsilon}),$$

where ε_{it} is i.i.d. over time, but is allowed to be contemporaneously correlated.

In a study of crime and education, Merlo and Wolpin (2008) find that there is an important role for observed and unobserved heterogeneity in initial conditions, as well as for stochastic events in determining outcomes as young adults. To consider the effect of unobserved heterogeneity, we introduce a finite number of individual types. To consider the effect of stochastic events, we introduce the four stochastic shocks included in ε_{it} . Finally, to capture the possibility that some individuals are more prone to consuming drugs based on observable initial conditions, the mean of ε_{it}^{sd} is specified as follows (an identical treatment is given to ε_{it}^{hd}):

$$\mu_{\varepsilon_{it}^{sd}} = \rho_1 MEd + \rho_2 HH + \rho_3 R + \rho_4 Ed13$$

where ρ_1, ρ_2, ρ_3, and ρ_4 are parameters to be estimated.

The State Vector

The state vector z_{it} consists of the following elements: the year in which the person was born; the years of completed education; schooling decision in the previous year; job experience; the type of household at age 12 (single-parent vs. two-parent household); mother's education; the past four periods of soft and hard drug consumption; accumulated wealth; non-earned income; current prices for soft and hard drugs; race; and years of education at age 13.[17]

[16] See Appendix A to see the functional form used to introduced unobserved heterogeneity in the model.

[17] Though the theoretical model described above considers all the past periods up until age 13, for practical reasons we limited the number of state variables up to 4 periods previous to the current one. This limitation does not affect in any significant way.

3.2 Adaptation of Bernheim-Rangel's model

As in Becker and Murphy's model, the individual is initially drawn to an addictive substance because it gives utility through the experience of a "high". However, once the person has consumed the drug, he can operate in two modes in each period: a cold mode in which he rationally chooses the best decision among all alternatives, and a hot mode in which he invariably consumes the addictive substance irrespective of his underlying preferences. The likelihood of entering a hot mode is assumed to increase with the level of past consumption to reflect the higher probability of encountering cues that trigger the hot mode.

To adapt Bernheim and Rangel's model to the case under study, we parameterize the probability of entering the hot mode for drug v as:

$$p_v(x) = \frac{\exp(\omega_0 + \omega_1 x_v + \omega_2 x_v^2)}{1 + \exp(\omega_0 + \omega_1 x_v + \omega_2 x_v^2)} \text{ if } x_v>0, \text{ and } p_v(x) = 0 \text{ otherwise}$$

where v=sd, hd, and x represent the history of use for drug v. Following Bernheim and Rangel, $x_v = 0, 1, ..., X_v$, and it evolves as follows: usage in state $x_v \geq 1$ leads to state $min\{X_v, x_v + 1\}$ in the next period, and no use leads to state $max\{1, x_v - 1\}$. State $x_v = 0$ represents the state in which the individual has had no experience with drug v, and it is unreachable once the drug has been consumed. Thus, the individual cannot enter a hot mode if he has had no contact with the drug before, but will be exposed to hot modes after consumption with positive probability. Environmental cues may trigger the hot mode, but only for drugs that a person has experience with. This assumption is reasonable because neurological and psychological analyses show that experience with a substance is necessary to be affected by cues that trigger the hot mode.

The maximization problem faced by the individual depends on whether the individual has had experience with drugs. If the person has not consumed any type of drugs in the past $p_{sd}(x_{sd}) = p_{hd}(x_{hd}) = 0$, and the expected payoffs can be represented as:

$$V_{S,W,Sav,a_t^{sd}=0,a_t^{hd}=0}(y_{it}, x_{sd} = 0, x_{hd} = 0, t) = U_{S,W,Sav,a_t^{sd}=0,a_t^{hd}=0}(y_{it}, 0, 0, t)$$

$$+\delta E[V(y_{t+1}, 0, 0, t + 1|y_{it}, 0, 0, d_t = S, W, Sav, a_t^{sd} = 0, a_t^{hd} = 0]$$

where $V_{S,W,Sav,a_t^{sd}=0,a_t^{hd}=0}$ is the alternative-specific value function that corresponds to a certain level of schooling, of working and of savings and no consumption of drugs ($a_t^{sd} = 0$ and $a_t^{hd} = 0$); y_{it} is the state vector for individual i in time t without including the history of soft and hard drug use (x_{sd} and x_{hd}, respectively); d_t represents the choice in t; and δ represents the discount factor. To simplify notation, we will denote from now on $V_{S,W,Sav,a_t^{sd}=0,a_t^{hd}=0}(y_{it}, x_{sd} = 0, x_{hd} = 0, t)$ as $V_{a_t^{sd}=0,a_t^{hd}=0}(x_{sd} = 0, x_{hd} = 0)$, $U_{S,W,Sav,a_t^{sd}=0,a_t^{hd}=0}(y_{it}, 0, 0, t)$ as $U_{a_t^{sd}=0,a_t^{hd}=0}$, and $E[V(y_{t+1}, 0, 0, t+1|y_{it}, 0, 0, d_t = S, W, Sav, a_t^{sd} = 0, a_t^{hd} = 0]$ as $E[V(0, 0|0, 0, d_t]$.

Similarly, if the person decides to consume soft drugs in the present, the expected payoff is:

$$V_{a_t^{sd}>0,a_t^{hd}=0}(x_{sd} = 0, x_{hd} = 0) = U_{a_t^{sd}>0,a_t^{hd}=0}$$
$$+\delta E[V(min\{X_{sd}, x_{sd} + 1\}, 0|0, 0, d_t]$$

The same logic follows for cases where $a_t^{sd} = 0$ and $a_t^{hd} > 0$ and $a_t^{sd} > 0$ and $a_t^{hd} > 0$.

In the case that a person has experience with both types of drugs and decides to consume both types of drugs in period t, the expected payoff is:

$$V_{a_t^{sd}>0,a_t^{hd}>0}(x_{sd}>0,x_{hd}>0) = U_{a_t^{sd}>0,a_t^{hd}>0}$$
$$+\delta E[V(min\{X_{sd},x_{sd}+1\},min\{X_{hd},x_{hd}+1\}|x_{sd},x_{hd},d_t]$$

However, if he decides not to consume hard drugs during the period, the individual knows that he is susceptible to cue-triggered mistakes, since $p_{sd}(x_{sd})>0$ and $p_{hd}(x_{hd})>0$ when one has been exposed to addictive substances. The individual will incorporate these probabilities in the calculations since he recognizes that he is susceptible to cue-triggered mistakes. If he rationally decides to consume soft drugs but not hard drugs, the expected payoff is now defined as:

$$V_{a_t^{sd}>0,a_t^{hd}=0}(x_{sd}>0,x_{hd}>0) = (1-p_{hd})U_{a_t^{sd}>0,a_t^{hd}=0} + p_{hd}U_{a_t^{sd}>0,a_t^{hd}>0}$$
$$+(1-p_{hd})\delta E[V(min\{X_{sd},x_{sd}+1\},max\{1,x_{hd}-1\}|x_{sd},x_{hd},d_t]$$
$$+p_{hd}\delta E[V(min\{X_{sd},x_{sd}+1\},min\{X_{hd},x_{hd}+1\}|x_{sd},x_{hd},d_t]$$

and the same logic follows for the cases where $a_t^{sd}=0,a_t^{hd}>0$ and $a_t^{sd}=0,a_t^{hd}=0$.

A nice feature of this model is that Becker and Murphy's model is a specific case of it. This occurs when $p_{sd}(x_{sd})=p_{hd}(x_{hd})=0$ for all x_{sd},x_{hd}, that is, when individuals do not make mistakes.

3.3 Solution Method

Recasting the maximization problem into a dynamic programming framework:

$$V(z_{it},t) = \max\{V_k(z_{it},t)\}$$

where the value function, $V(z_{it},t)$, is equal to the maximum over the alternative-specific value functions for each choice, V_k, with k=1,...,K. The value of choosing any of these k mutually exclusive choices is:

$$V_k(z_{it},t) = U_k(z_{it},t) + \delta E[V(z_{it+1},t+1|z_{it},d_t=k]$$

where δ denotes the discount factor.

To calculate expectations we use Monte Carlo integration using an antithetic estimator to integrate over the shocks $(\varepsilon_{it}^w,\varepsilon_{it}^s,\varepsilon_{it}^{sd},\varepsilon_{it}^{hd})$. Individuals predict their non-earned income and the prices of drugs for the following period to calculate expectations. To do this, non-earned income is discretized and we assumed that individuals know the transition probability of having a certain level of non-earned income in period t+1 given the level in period t.[18] We also assume that prices for drugs evolve deterministically and that people have correct point expectations about future prices.[19] Individuals also consider the probability of repeating a grade if attending school.

[18] Transition probabilities are calculated from the data. For more details, see the Data section.

[19] Arcidiacono et al. (2007) consider the same simplification about expectations of prices in the context of alcohol and cigarette consumption.

Given the finite horizon, the solution method proceeds by backward recursion. Since the state space is too large to calculate the exact value function for each possible state, we calculate the value function for each age for a randomly chosen subset of states. We use these values to run a polynomial regression on the state space to approximate all the value functions for all the states for each age. Keane and Wolpin (1994) showed that this method works well in approximating the full solution of a dynamic programming problem. To avoid the computational burden of solving the model until the last period, we begin the backward recursion at age 45, using the polynomial form of the value function at that age as the terminal condition. The parameters associated with this terminal condition are jointly estimated along with the structural parameters of the model.[20]

3.4 Estimation Method

To estimate the model we follow the estimation technique introduced in Keane and Wolpin (2001) and Keane and Sauer (2010) and applied in Keane and Sauer (2009) and Buchinsky and Gotlibovski (2008). This technique is based on simulated maximum likelihood, allowing for classification errors in discrete choices and measurement error in continuous outcomes. The incorporation of classification error has three major advantages. First, the zero probability problem often faced in a pure frequency simulation method is avoided given that there is a positive probability that any simulated history is the individual's true choice history with classification error. Second, choice probabilities are computed from unconditional simulations of the model, rather than being conditional on the state space, addressing the problem of unobserved state variables. These advantages are relevant here because the choice set and the state space are both large. Data constraints do not allow us to observe all of the choices and state variables for every period the individual is followed.[21] Finally, this methodology is based on the realistic assumption that datasets contain errors and proposes a way to clean estimates from biases introduced by these errors under certain assumptions.[22]

Following Buchinsky and Gotlibovski's (2008) notation, we define the data consisting of outcome histories as $\{C_i^*, w_i^*, x_i\}_{i=1}^N$, where $C_i^* = \{c_{it}^*\}_{t=13}^T$ is the history of reported choices (decisions on attending school, working, consuming drugs, and holding assets), $w_i^* = \{w_{it}^*\}_{t=13}^T$ is the history of reported wages, x_i is a vector of initial conditions for individual i (including race, mother's education, years of completed education at age 13, non-earned income at age 13, and type of household at age 12), N denotes the sample size, and T denotes the last age at which individuals are observed (25 years old in this case).

To account for missing initial conditions, choices, and accepted wages, we define three additional variables: $I(x_i^*) \equiv I(x_i^* \text{ is observed})$, $I(c_{it}^*) \equiv I(c_{it}^* \text{ is observed})$, $I(w_{it}^*) \equiv I(w_{it}^* \text{ is}$

[20] The exact specification of the terminal condition and the polynom used to approximate the value functions for the rest of the periods can be found in Appendix A: Functional Forms.

[21] Wealth, for example, is observed only in some interviews. For more details, see the Data section.

[22] For example, Poterba and Summers (1986) and Keane and Sauer (2009) note the problem of misclassification of work status. The authors conclude that failing to control for misclassification can lead to important biases. Buchinsky and Gotlibovski (2008) arrived at the same conclusion when considering misreporting from retrospective data. Bound and Krueger (1991) and Gottchalk and Moffitt (2008) consider the effect of measurement error in wages.

observed). These are indicator variables that take the value of one if $I(.)$ is true and zero otherwise.

The algorithm proceeds as follows:

Step 1. Given a set of parameters, draw initial conditions, x_i, for M artificial individuals of type g.

Step 2. Draw $(\varepsilon_{it}^{w}, \varepsilon_{it}^{s}, \varepsilon_{it}^{sd}, \varepsilon_{it}^{hd})$ for each M. Construct the alternative-specific value functions $V^k(\varepsilon_{it}^{w}, \varepsilon_{it}^{s}, \varepsilon_{it}^{sd}, \varepsilon_{it}^{hd})$ for period one, for k=1,...,K (where K is the total number of choices) and choose the c_{i1}^{*} along with the corresponding w_{i1}^{*} that maximizes utility.

Step 3. Update the state variables.

Step 4. Go to t=2. Repeat steps (2) and (3).

Step 5. Go to t=3. Repeat steps (2) and (3), etc., until T is reached.

After generating M simulated outcome histories $\{C_i^*, w_i^*, x_i\}_{i=1}^{M}$ for each type g:

Step 6 Compare x_i for the first observed individual with x_i for the M simulated individuals to compute classification error rates π_{jlx_i}, to be defined below, which allow the probability of reporting a particular initial condition to differ from the true one. That is, compute $\widehat{\pi}_{jlx_i} = \Pr(\text{reported } x_i = j | \text{actual } x_i = l)$, for values of l and j that depend on the specific initial condition. For example, for mother's education l and j take values of 0 and 1, representing complete high school education or less, and at least some college.

Step 7 Compare c_{it}^{*} for the first observed individual with c_{it}^{*} for the M simulated individuals to compute classification error rates π_{jlt} for each choice contained in c_{it}^{*}.

Step 8 Form the type-specific likelihood contribution for the first individual as:

$$\widehat{P}_g(C_i^*, w_i^*, x_i|\theta) = \frac{1}{M}\sum_{m=1}^{M}(\sum_{j=j_0}^{J}\sum_{l=l_0}^{L}\widehat{\pi}_{jlx_i}[x_i = j, x_i^m = l])^{I(x_i^*)}$$

$$\prod_{t=1}^{T}(\sum_{j=j_0}^{J}\sum_{l=l_0}^{L}\widehat{\pi}_{jlt}[c_{it}^* = j, c_{it}^m = l])^{I(c_{it}^*)} f_w(\varepsilon_{it}^m)^{I(w_{it}^*)}$$

where θ is the vector containing all the model parameters, and $f_w(\varepsilon_{it}^m)$ is the measurement error density in reported wages.

Step 9. Repeat steps 6 to 8 for the rest of the N observed individuals.

Step 10 Repeat steps 1 to 9 for each type g to compute the likelihood contribution for each individual i:

$$\widehat{P}(C_i^*, w_i^*, x_i|\theta) = \sum_{g=0}^{G} \tau_{g|IC} \widehat{P}_g(C_i^*, w_i^*, x_i|\theta)$$

where $\tau_{g|IC}$ denotes the probability of being of type g, with g=1,...,G, conditional on the initial conditions.[23]

From this procedure, it is clear that the probabilities $\widehat{P}(C_i^*, w_i^*, x_i|\theta)$ do not depend on the state variables at any time t; they depend only on the outcomes of the model (initial conditions, choices and wages).

Classification Error Rates

All the discrete outcomes are assumed to be subject to classification error, that is, there is some probability that the individual's response is the truth and some probability that it is not. These classification errors are assumed to be unbiased.[24] This implies that the probability a person reports to choose an option is equal to the true probability that the person chooses that option. It also implies that the true classification error is linear in the true choice probability.[25]

As an example, consider the structure of the classification error for the working decision. Define $\pi_{jlac} = P(reported\ work_{ac} = j|actual\ work_{ac} = l)$ for j,l=0,1, where a denotes age and c denotes cohort.[26] Thus, for each age-cohort combination we will have four classification rates:

$$\begin{aligned}
\pi_{11} &= P(reported\ work = 1|actual\ work = 1) \\
\pi_{10} &= P(reported\ work = 1|actual\ work = 0) \\
\pi_{01} &= P(reported\ work = 0|actual\ work = 1) = 1 - \pi_{11} \\
\pi_{00} &= P(reported\ work = 0|actual\ work = 0) = 1 - \pi_{10}
\end{aligned}$$

with

$$\begin{aligned}
\pi_{11} &= CE_w + (1 - CE_w) f(work_{ac} = 1) \\
\pi_{10} &= (1 - CE_w) f(work_{ac} = 1)
\end{aligned}$$

where $f(work_{ac} = 1) = \frac{1}{M} \sum_{i=1}^{M} I(work_{ac} = 1)$, $I(.)$ is an indicator function that takes value 1 when (.) is true, M is the number of simulated histories for each cohort and $0 < CE_w < 1$

[23] For the exact specification of $\tau_{g|IC}$, see Appendix A: Functional Forms.

[24] This methodology can be easily extended to allow persistence in misreporting.

[25] For a more detailed explanation about classification errors, identification and consistency of the method, see Keane and Sauer (2009).

[26] The introduction of a cohort is necessary because people born in different years face different prices of drugs given their age. If prices of drugs affect drug behavior and drug behavior affects other behaviors, aggregating simulated histories for a certain outcome for a certain age to calculate the average behavior would be incorrect.

is a parameter to be estimated. Since π_{11} approximates CE_w when $f(work_{ac} = 1)$ approximates 0, CE_w can be interpreted as the probability that low probability events are classified correctly.

The same procedure is used to define classification error rates for the other discrete outcomes in the model: school attendance, consuming soft and hard drugs, mother's education, level of education at age 13, race, household type at age 13, non-earned income received at age 13, and savings. Although saving is a continuous variable, in this analysis it has been discretized in four categories, so a classification error approach has been used for this variable.

Measurement Error

Wages are assumed to be subject to classical measurement error. This assumption allows us to reconcile the history of simulated wages with reported ones. Specifically,

$$w_{it}^{obs} = w_{it} \exp(\varepsilon_{it}^{m})$$

where w_{it}^{obs} is the hourly wage reported by individual i in period t, w_{it} is defined as in Equation (3), and ε_{it}^{m} is the wage measurement error, which is assumed to be distributed as follows:

$$\varepsilon_{it}^{m} \sim N(0, \sigma_m^2).$$

Thus, if the individual reports to work and reports a wage associated with that job, the likelihood is multiplied by the density function of the measurement error if the simulated history reflects that the person works. This density is:

$$f(\varepsilon_t^{w,m}; \sigma_m^2) = \frac{1}{\sigma_{w,m}} \phi(\ln(\frac{w_t^{obs}/w_t}{\sigma_m}))$$

where ϕ denotes the density function of a standard normal variable and σ_m is a parameter to be estimated. If the wage is not observed, the likelihood is simply multiplied by $f(\varepsilon_t^{w,m}; \sigma_m^2) = 1$.

4 The Data

The data for this study are taken from the National Longitudinal Survey of Youth 1997 (NLSY97). The NLSY97 consists of 8,984 respondents, 4,599 of them men, who were born during the years 1980 through 1984. The sample consists of a core random sample representing the U.S. population born in the years 1980 through 1984 and supplemental samples of blacks and Hispanics. The first interviews were conducted in 1997, with subsequent interviews conducted annually thereafter. The analysis here is based on 4,423 males who were interviewed at least three times from 1997 to 2006.

The NLSY97 collects schooling and employment data in event history form, as well as information in each survey round about frequency of consumption of marijuana and other drugs, also referred as soft and hard drugs in the questionnaire.

The discrete decision period is assumed to be a school year; that is, the period that starts on October 1 and ends on September 30 of the following year. Below, we define the specific rules we used to classify annual activity for each individual.

4.1 Definition of Choices and Variables

In this subsection we define the main variables of the analysis. For information on other variables used in the analysis and details about how we create a price index for hard drugs, see Appendix C: More on Definition of Choices and Variables.

Schooling

To determine school attendance, we look at an individual's activity each month of every calendar year, barring the period from July through September. If a student attended school for at least six months or if the individual reported completing an additional grade level by October 1 of the next year.[27]

Working

To assign annual work status, we use the NLSY97 work histories for all weeks between October 1 and June 30 of each year. Note that the summer quarter is not taken into account to avoid categorizing someone with a summer job as working. An individual is considered to have worked during the year if he was employed for at least two-thirds of the weeks. If work status is missing for less than two thirds of the weeks, then we apply the same criterion as before for the weeks with information. If more than two thirds of the weeks have missing information, the variable is classified as missing for the period. If the person is classified as working during the period, it is assumed that he works 2000 hours during the year.

Real Wages

Nominal hourly wages are defined using hourly wages for the main occupation. These nominal wages are translated into real wages by dividing by the Consumer Price Index expressed in 1983 dollars.

Drugs

Starting in the first round of the survey, respondents were asked if they have ever used marijuana. They were asked at what age they first tried marijuana and the frequency (in days) of consumption over the previous 30 days. All of this information was used to determine if and when a person consumed marijuana.

If the person responds affirmatively, the amount consumed is assumed to be equal to the median frequency of consumption observed in the data. This amount is multiplied by twelve

[27] The schooling data were hand-edited in order to obtain a consistent profile of attendance and grade completion. I do this using other information provided in the NLSY97, that is, the monthly attendance calendar, highest grade attended, highest grade completed, history of grade attended and school progression by month and year, history of college enrollment, highest grade completed before the new school year, and date when high school diploma was received (excluding GEDs). If attendance was not possible to determine, the variable was coded as missing.

to express it in annual terms. It is also assumed that in a day a person consumes half a gram of marijuana, which is approximately the amount in a marijuana cigarette.

Beginning in the second round of the survey, respondents were asked about the use of any drugs such as cocaine, crack, heroin, or other substances not prescribed by a doctor (also referred in the questionnaire as hard drugs) in order to get high or to achieve an altered state. Details collected include the respondent's age at first use, and the number of times since the last interview that the respondent took some drug or other substance. The amount consumed is given by the median frequency of consumption observed in the data.

Savings

Respondents were asked about asset and debt holdings at the time they turned 18, or became independent, and once again in the first interview after they turned 20 and 25.[28] Savings are expressed in real terms using the 1983 Consumer Price Index. The level of assets was discretized to four groups. The categories are given by: (a) less than \$-1,000, (b) between \$-1,000-\$3,000, (c) \$3,000-\$20,000, and (d) more than \$20,000. The value assigned to each category differs by age, and is predicted by the following quantile regression:

$$Assets_{ig} = \alpha_0 + \alpha_1 Age_{ig} + \alpha_2 Age_{ig}^2 + \upsilon_{ig}$$

where the sub-index ig denotes the individual i who belongs to group g, with g=1,2,3,4, and υ_{ig} is assumed to be i.i.d.. A quantile regression was used rather than an OLS regression to avoid the effect of outliers on the prediction of wealth.

Prices of Drugs

The Drug Enforcement Administration (DEA) has been collecting information on drug transactions made by federal agents, law enforcement officers and their informants for the System to Retrieve Information from Drug Evidence (STRIDE) database. Even though some analysts have raised questions about the validity of the STRIDE data as a source of prices of drugs (Horowitz 2001, Manski et al. 2001), in a recent study, Arkes et al. (2008) demonstrate that when substantive knowledge about drug markets is incorporated into the use of STRIDE data, the data do in fact generate consistent trends across agencies within the same market. They conclude that STRIDE is useful for estimating trends in the price of illicit drugs. The price of marijuana used in this analysis was taken from the website www.drugscience.org. It is an annual average price for one gram of marijuana based on purchases of 10 to 100 grams. Prices for the period 1993-2003 were constructed using the STRIDE database, and for the period 2004-2005 using the National Survey on Drug Use and Health. To construct a price index for hard drugs, we use prices reported by Arkes et al. (2008) and information reported by Agar et al. (1998). For details about its construction, see Appendix C.

4.2 Descriptive Statistics and Observed Patterns

Table B.1 (see Appendix B for tables and graphs) presents summary statistics of the variables used in the analysis. On average, 22 percent of respondents in the sample consumed soft

[28] NLSY97 considers an individual independent when he/she turns 18, moves out from his/her parents' house, gets married or has a child.

drugs each period, while 6 percent consumed hard drugs; 50 percent were white; 52 percent lived in a household where both parents (biological or not) were present at age 12; 68 percent were in grade 7 or 8 at age 13, while the remaining were in grades 5 or 6. This means that 32 percent of respondents in the sample were behind in school by age 13; and 41 percent of the sample belonged to a house where the mother had at least some college education.

In the analysis, we make the assumption that individuals start deciding whether to consume drugs at age 13, implying that they are free from the effect of drugs when they enter the sample. Table B.2 shows the age of initiation to both types of drugs. Only 6 percent of the sample consumed soft drugs before age 13 (column 4), and 1.78 percent consumed hard drugs (column 7), so the assumption seems realistic. Table B.3 shows the distribution of drug consumption by age. It suggests that the initiation to drugs happens approximately between 13 and 20 years of age.

Column 2 shows that the distribution of non-users by age is U-shaped, with the minimum at the age of 18, when the number of people who start experimenting with drugs falls below the number of people who stop consuming.

It is apparent from table B.3 that people who consume hard drugs usually consume soft drugs at the same time (columns 3, 4, and 5). This is even more striking when one compares people who have never consumed drugs to people who have consumed only soft drugs, only hard drugs, or both for the whole period. According to this classification, only 1.8 percent (74 individuals) have consumed hard drugs without ever consuming soft drugs (37.7 percent have not consumed drugs, 37.4 percent have consumed only soft drugs, and 23.1 percent have tried soft and hard drugs).

According to the model, a person is affected by drug consumption if he consumed drugs in the current period and/or any previous period. However, to keep the number of state variables in a reasonable number due to computational constraints, in practice we keep track of the last four periods of drug consumption. This implies that after five periods of no consumption, the person is cleaned again from any side effect of drugs. Given this assumption, we define a person to be a drug user if he has consumed soft or hard drugs in the current or in any of the last four periods, and non-drug user otherwise. Though this assumption could seem restrictive, the following patterns are not changed when more lags are considered to define who can be affected by the effect of drugs.

One striking difference between drug users and non-drug users is the number of completed years of education and school attendance. Table B.4 shows that even though everyone starts with the same number of years of education, by the age of 24 non-drug users have 0.75 and 0.92 more years of education on average than marijuana and hard drugs users, respectively. Attendance follows a similar pattern (Table B.5), so that the difference rapidly increases at age 16, and reaches its maximum between ages 17 and 18, when the attendance rate between non-drug users and marijuana users differs by 18.5 percent and by 22.5 percent between non-drug users and hard drug users.

There are several explanations for these patterns. For example, lower returns to education for those who consume drugs might create a disincentive to invest in education; drug users may need money to buy drugs and be less likely to enroll in college education to avoid its monetary cost; unobserved heterogeneity or observed initial conditions could also contribute to the empirical pattern. The structure of the model presented in section 3 separates the channels for these observed patterns.

Graph B.3 shows that after age 16, non-drug users tend to work more than marijuana users and other drugs users, with a participation rate almost 10 percentage points higher for non-drug users at age 24. Graph B.4 shows that there is no clear pattern for hourly wages. Wages for non-drug users are the highest except for the age range 16-17, though they are very similar to those of hard drug users. For marijuana users, wages are similar to those of the rest of the groups, but at age 21 they start diverging, showing a slower rate of growth.

Finally, graphs B.5 to B.12 show the percentage of marijuana users and hard drug users according to the observed initial conditions of the model, and the patterns differ for all of the initial conditions. The participation in marijuana consumption is very similar at younger ages for whites and non-whites, but consumption is higher among whites after age 15 (Graph B.5). On the other hand, the participation in consumption of hard drugs is always higher among whites (Graph B.6). Marijuana consumption is also higher among those living in a single-parent household at age 12 (Graph B.7), while hard drug consumption seems not to differ by the type of household the child was raised in (Graph B.8). In terms of consumption by the mother's level of education, the patterns seem not to differ at younger ages, and is higher for both drugs after age 18 for those with more educated mothers (Graphs B.9 and B.10). Participation in consumption for both drugs also differs according to level of education at age 13. Those with less education at the beginning of the analysis (five years of education) consume more than others (Graphs B.11 and B.12)

5 Estimation Results

The results in this paper are based on estimating the simplest version of the model presented in Section 3, assuming that individuals are rational and time consistent, as in Becker and Murphy's model of rational addiction. It does not contain unobserved heterogeneity either. In future work we will incorporate unobserved heterogeneity by the inclusion of a finite number of types of individuals and allowing addicts to commit mistakes, as in Bernheim and Rangel's model.

5.1 Parameter Estimates

Estimates of the parameters are provided in Tables B.6 through B.11 in Appendix B together with standard errors.[29] In Table B.6 we present the parameter estimates associated with the utility function. In Table B.7 we present the wage equation and school related parameter estimates. In Table B.8 we present the parameter estimates of the variance-covariance matrix. In Table B.9 we present the parameters associated with the initial conditions. In Table B.10 we present the parameter estimates of the classification and measurement errors. In Table B.11 we present the parameter estimates of the probability of repeating a grade. Finally, in Table B.12 we present the parameters of the terminal value function.

Given the CRRA utility function for consumption of a composite good (Table B.6), $\frac{c_{it}^{\alpha_c}}{\alpha_c}$, the estimate of α_c implies an intertemporal elasticity of substitution (IES) of consumption

[29] Some of the standard errors seem too small to be correctly estimated. This can happen if the likelihood function is discontinuous around the parameter estimates. In a future version, we will present the standard errors calculated by bootstrap.

of -2.29, similar to the estimated value of -2.07 obtained by Keane and Wolpin (2001). In the case of leisure, the estimated IES is approximately -4.25.

According to the estimates of β_{sd} and β_{hd}, the coefficients that measure the importance of past consumption in the utility function for drugs, past consumption of soft and hard drugs matter for the decision of current consumption. However, the depreciation rates are very different: while marijuana effects depreciate fast over time ($\beta_{sd} \simeq 0.11$), the reverse occurs with hard drugs ($\beta_{hd} \simeq 0.63$).

Estimates of the wage equation are presented in Table B.7. The following example will provide intuition on the meaning of these coefficients. In Table B.13 we calculate hourly wages under different scenarios for the average person in the sample. For this example, let us consider a person who has not consumed drugs in the current or previous four periods, that is, a non-user. The first measure, the returns to school for a non-user, is 9.98 percent (column 2). This is, an additional year of schooling increases the hourly wage in 9.98 percent for a non-user. If the person has consumed marijuana today and in the previous four periods in a low dosage (12 days each year), his return to education decreases slightly to 9.92 percent (column 3). This figure drops further to 9.39 percent if the person increases his consumption to 120 days per year (column 4), and is as low as 8.49 for someone who consumes marijuana almost daily (300 days each year, column 5). If the average person consumes only hard drugs in low dosages today and in the previous four periods (8 times each year, which is the observed median value of hard drugs consumed in the data), his return to school is 9.45 percent (column 6). If the person decides to consume hard drugs as often as once a week (52 times each year, which is a measure used in the literature to distinguish chronic or problematic hard drug users from occasional or non-problematic users) his return to education goes down more than 3 percentage points, to 6.51 percent (column 7). Hence, drugs have an important effect on returns to school when consumed in high dosages. The estimates also suggest that hard drugs have a more negative impact than marijuana, and that marijuana, consumed in small dosages, has very little impact on the returns to school.

The second measure presented in Table B.13 is the return to work experience for the average person in the sample for different levels of drug consumption. Returns to work experience also decrease with the dosage of marijuana and hard drugs consumed, but are similar to the returns for a person who has not consumed drugs today or in the previous four periods. Thus the effects are much less pronounced than those observed for education.

Table B.14 compares the hourly wage rate for people with varying levels of education and work experience, across different levels of marijuana and hard drug consumption. Hourly wages for a person who consumes low levels of marijuana (12 days per year) are slightly lower than the hourly wages for a person who abstains. This difference increases with the level of marijuana consumption and the level of schooling, given the negative effects that marijuana has on the returns to school. Hourly wages follow the same pattern if the person decides to consume hard drugs instead. However, the reduction is more pronounced for hard drug consumption than for marijuana consumption for similar levels of consumption.

The estimates of the probability of not advancing in school (Table B.11) present the expected signs. Working while studying increases the probability of not advancing in school, while having attended school the previous period reduces it. Individuals with more years of completed education are less likely to repeat a grade, but older individuals have more chances of not advancing. Individuals that are white, whose mothers have at least college

education, who lived in a household with two parents at age 12 and that were not behind in school at age 13 are less likely to repeat a grade. Finally, soft and hard drug consumption, in the present or in the past, increases the likelihood of not advancing in school.

Correlation coefficients among errors (Table B.8) suggest that schooling and working decisions are negatively correlated with consumption of marijuana and positively correlated with consumption of hard drugs. This implies that those who face higher costs of schooling are less (more) likely to consume marijuana (hard drugs), and those with higher draws for wages will be less (more) likely to consume marijuana (hard drugs). Also, those who face higher draws for wages are more likely to face a lower cost of schooling; and the stochastic shocks to taste of drugs are highly and positively correlated, implying that those who like one type of drug are more prone to like the other types too. All these correlation coefficients are statistically significant, which offers evidence in favor of the importance of acknowledging the contemporaneous simultaneity of these decisions.

The estimates for the initial conditions (Table B.9) show that some of the factors that make individuals observably heterogeneous in the initial period make some individuals more or less prone to consume drugs than others. For example, those with 6 years of completed education at age 13 are more prone to consume marijuana than those with 5 years of education. The contrary occurs to those that were living with both parents at age 12, those with a mother with at least college education, and those with 8 years of education at age 13. For the case of hard drugs, only having 6 years of completed schooling vs. 5 years, and a mother with at least some college make individuals more prone to like hard drugs. The rest of the variables are not statistically significant.

5.2 Model Fit

Comparisons to gauge the fit of the model provide a sense of the overall credibility of the model. Data simulated from the model differ from actual data. However, the model fits the overall patterns in the data quite well, as shown below. This provides confidence in the ability of the model to analyze counterfactual exercises.

To examine the goodness of fit of the model we compare its predictions with observed counterparts in the NLSY97. To do so, we split the sample into non-drug users, marijuana users and hard drug users. In the model only the current and previous four periods of consumption of a drug are relevant for deriving utility from drugs and for the effect of drugs on wages. Thus, we define a person as a non-drug user if he has not consumed any type of drug in the current period or in the four previous periods. Similarly, a person who has consumed marijuana (hard drugs) in the current period and/or in any of the previous four periods is classified as a marijuana (hard drugs) user.

Figures B.13 and B.14 offer a comparison of the actual and predicted values for school attendance. In the NLSY97, non-drug users have the highest attendance rates for the age range observed, while hard drug users have the lowest attendance rates. Figure B.14 shows that the model captures very well the pattern of school attendance for the sample as a whole. The model is also able to capture some of the differences in attendance after age 21 by type of user. However, the model does not predict the differences in attendance rates observed in the NLSY97 during early ages, overpredicting the attendance rates for marijuana and hard drug users.

In Figures B.3 and B.15 we compare the actual and predicted value for the employment rate. In the NLSY97, the employment rate of non-drug users and drug users are very similar until age 19. After this, the employment rate for non-drug users grows faster than the employment rate for marijuana and hard drug users. Figures B.15 shows that the model captures patterns observed in the data very well. Although the model predicts lower employment rates at ages 16-17 and 21-24 than the observed ones in actual data, the model is able to capture the similarity in employment rates until age 19 for all types of users, and the higher employment rate for non-drug users with respect to drug users after age 19.

In Figures B.4 and B.16 we compare the actual and predicted value for hourly wages. In the NLSY97, wages among users and non-users are similar until age 20. Afterwards, the hourly wage for non-drug users is higher than the hourly wage for marijuana users, which grows at a slower rate. The wage for hard drug users fluctuates between the values for the other two groups. Figure B.16 shows that the model predicts higher wages for non-drug users, and the difference in wages becomes more pronounced with age when comparing non-drug users with marijuana users. In the case of hard drug users, the model predicts that their wages are lower than the observed ones.

Finally, in Figures B.17 and B.18 we compare the actual and predicted percentage of men who consume drugs by age. Even though the model overstates marijuana consumption at early ages and understates it for ages 15 to 22, it accurately captures the bell-shaped pattern observed in the data. In the case of hard drugs, the model predicts a more erratic pattern than the one observed in the data, with higher consumption for ages 13 to 15, but relatively close to the observed values for the rest of the period.

5.3 Results from Experiments

In this section we report results from two different experiments. The first experiment analyzes the effect of a change in the price of marijuana. The US government spends most of its drug control budget on reducing the supply of drugs thereby raising their prices.[30] In this experiment, we consider a permanent rise in the price of marijuana consisting of a 30 percent increase in the price each year. This experiment would shed light on the effect that those price changes due to government intervention have on individuals' education and working decisions, wages and drug consumption. The second experiment examines the question of whether consuming drugs in early stages of life (adolescence) has detrimental effects on labor market outcomes and educational attainment. To do this, we simulate the choice paths for 100,000 individuals in two alternative situations: one in which they can consume drugs at any age (status quo) and another one where they can only consume drugs after adolescence (age 19). Comparing these two scenarios, we are able to determine if drugs consumed in early ages have a permanent effect in latter outcomes.

[30] According to the US office of National Drug Control Policy (ONDCP), federal spending on illegal drug control was $19 billion in 2002. 67 percent of this expenditure was directed at reducing the supply of drugs through spending on domestic law enforcement (50 percent), interdiction (11 percent) and international activities (6 percent).

5.3.1 Do government policies to affect prices of drugs have an impact on labor market outcomes, educational attainment and drug consumption?

The experiment on increasing the price of marijuana by 30 percent for the period of analysis suggests that governments can affect certain behaviors and outcomes with policies that affect the supply of drugs. According to this experiment, an increase of 30 percent in the price of marijuana decreases the number of people that consumes marijuana by 23 percent on average (Figure B.19). This implies that the demand for marijuana participation is inelastic, with elasticities that range from 0.69 at age 13, to 0.82 at age 25. An increase in price would imply a decrease of as much as 5.2 percentage points for the age range in which consumption is highest (18 to 22 year-olds). Figure B.20 shows that despite the drop in marijuana consumption and the high correlation between taste shocks for these drugs, hard drug consumption is not altered by changes in the price of marijuana. Though this result may suggest that marijuana does not work as a drug of initiation, or as a gateway to using more dangerous drugs (gateway drug theory), the model does not explicitly incorporate the effect of past marijuana consumption on the current utility of consuming hard drugs.

Figures B.21 and B.22 present the effect of a reduction in price on school attendance and on the highest grade completed for the subset of the population that consumed marijuana with the low set of prices but was dissuaded from consuming marijuana with the higher set of prices. According to the simulation, the attendance rate of this group increases between 2 and 2.5 percentage points for the age range 19-25. These changes are statistically significant at the 10 percent level for ages 19-20 and at the 5 percent level for ages 21-25. Higher attendance results in 0.26 more years of completed education by age 29 for this group. A smaller increase in attendance rates is also observed for individuals who continue consuming marijuana under the new set of prices (not shown).[31] Their attendance rate increases 1 percentage point for the age range 19-25, and this difference is significant at the 5 percent level. This increase in attendance results in 0.08 more years of completed education by age 29 for this group, but this small difference is not significant at the 10 percent significance level. This group of consumers also reduces the frequency of marijuana consumption from 5.04 to 4.74 periods of consumption in response to the price increase. Finally, the attendance rate for the group that does not consume marijuana before and after the price change is not affected (not shown).

Figures B.23, B.24, B.25 and B.26 present the effects of an increase in the price of marijuana in labor market outcomes. Figure B.23 shows that the employment rate for those who stopped consuming marijuana due to the price increase rise by 2 percentage points on average, with rates ranging from 1.6 to 2.6 points. This difference is significant at the 5 percent level for the age range 22-29. Hourly wages for this group also increase (Figure B.24). This increase implies an hourly wage 63 cents higher at age 29, which represents $2,763 more dollars a year in 2010 dollars. Figure B.25 shows that the employment rate for those who consume marijuana before and after the price change also increases. This increment is around 1.5 percentage points for the group aged 19-29, and as much as 2.4 percentage points average for the group aged 16-19. The rise in the employment rate reflects the fact that people that consume marijuana under both set of prices consume marijuana less frequently, but also that consuming marijuana has become more expensive, so some individuals have to

[31] For results not shown in graphs or tables, contact alvaro.a.mezza@frb.gov

start working to have enough income to support their consumption. This liquidity constraint effect is more notorious among young people, where the highest increases in employment are observed. Figure B.26 shows that despite the effect on employment, hourly wages for those who continue consuming marijuana are not affected. Finally, those individuals who do not consume marijuana before and after the price change do not modify their labor choices, as expected (not shown).

A limitation of the analysis is that it does not capture the expected reduction in marijuana consumption levels per period for those who consume it before and after the price increase. This is because people decide whether to consume or not, but once they decided to consume, the amount consumed is determined by the median amount observed in the data. Thus, the results of this experiment should be interpreted as a lower bound of the total effect; the effects for those who consume marijuana before and after could be biased downward if the individuals are not able to reduce their consumption of marijuana with the increase in prices. Also, the effects on education are expected to be higher than reported here, given that the model is not able to predict all the observed differences in attendance rates between non drug users and drug users.

5.3.2 Does Consuming drugs during early stages of life affect later labor market outcomes and educational attainment?

To be done.

6 Conclusions

In this paper we extend the addiction models of Becker and Murphy (1988) and Bernheim and Rangel (2004) by explicitly introducing individuals' education and employment decisions within a dynamic finite horizon model of drug consumption choice. Incorporating joint decisions on education, employment and drug consumption in the setup allows me to empirically analyze the causal effect of illegal drug consumption on education and labor market outcomes, explicitly addressing the simultaneity and the dynamics that characterize these decisions.

According to preliminary estimates based on extending Becker and Murphy's model of addiction, the main effects of drugs acts through wages. Contrary to findings in the literature, marijuana and hard drug consumption negatively affect wages. This reduction is small for individuals who consume marijuana in very low doses, but high for individuals who consume marijuana on an almost daily basis (300 days a year). Hard drugs have a significant negative effect on wages even for individuals who consume them frequently in low dosages (eight times per year). This negative effect is very pronounced for chronic hard drug users.

A counterfactual experiment suggests that a 30 percent increase in the price of marijuana each period would reduce the number of marijuana consumers among 13-29-year-olds by 23 percent on average, implying an inelastic price elasticity of demand. Individuals who are dissuaded from consuming marijuana due to the higher price would have a higher labor income and would have slightly higher attendance and employment rates.

A limitation of the analysis is that results are based on the assumption that individuals

are rational and time consistent, as in Becker and Murphy's approach to addictions. Even though empirical evidence supports the rationality assumption, current research is not able to determine whether drug consumers are rational and time consistent or rational individuals who commit mistakes. In future work we will compare how Becker and Murphy's model of rational and time consistent agents performs with respect to Bernheim and Rangel's model of rational addicts who frequently commit mistakes. Theoretically, Bernheim and Rangel's model is supported by recent evidence from neuroscience, psychology and clinical practice. It is also able to account for patterns associated with addictions that other models, such as Becker and Murphy's model or Gruber and Koszegi's model of hyperbolic discounting, are not able to account for. It is of interest to study how much of a change these differences imply in explaining and predicting behavior. Different models of addictions have different implications in terms of welfare. While a corollary of Becker and Murphy's rational addiction model is that the government should not intervene in this market except when there are externalities, Bernheim and Rangel's model supports government intervention under certain circumstances. This is another reason to study which model fits the data the best.

Another limitation of this study is that the only type of available data on soft and hard drug consumption representative of US population is based on self-reports. It is expected that the figures are biased downwards. Also, this model studies the decision of whether to consume soft and/or hard drugs, but not the decision of how much to consume. Thus, experiments on changes in the price of drugs are informative about the number of people who start or stop consuming drugs, but not about drops or increases in quantities consumed due to the price change. Finally, the oldest people in the sample are 25 years old in the last round used in the analysis. If the effects of drugs take a long time to materialize, results presented in this study are expected to be lower bounds.

References

Agar M., Bourgois P., French J., Murdoch O. 1998. Heroin addict habit size in three cities: context and variation. *Journal of Drug Issues*, 28: 921-940.

Arcidiacono P., Sieg H., Sloan F. 2007. Living rationally under the volcano? An empirical analysis of heavy drinking and smoking. *International Economic Review*, 48: 37-65.

Akerlof G. 1991. Procastination and Obedience. *American Economic Review*, 81: 1-19.

Arkes J., Pacula R., Paddock S., Caulkins J., Reuter P. 2008. Why the DEA STRIDE data are still useful for understanding drug markets. *NBER Working Paper Series*, 14424.

Barnes G.E., Barnes M.D., Patton D. 2005. Prevalence and Predictors of "Heavy" Marijuana Use in a Canadian Sample. *Journal of Substance Use and Misuse*, 40: 1849-1863.

Becker G., Murphy K.M. 1988. A theory of rational addiction. *Journal of Political Economy*, 96: 675-700.

Bernheim D., Rangel A. 2004. Addiction and cue-triggered decision processes. *American Economic Review*, 94: 1558-1590.

Bernheim D., Rangel A. 2005. Behavioral public economics: welfare and policy analysis with non-standard decision makers. *NBER Working Paper Series*, 11518.

Bound J., Krueger A. 1991. The extent of measurement error in longitudinal earnings data: do two wrongs make a right?. *Journal of Labor Economics*, 9: 1–24.

Buchinsky M, Gotlibovsky C. 2008. Residential location, work location, and labor market outcomes of immigrants in Israel. *Working Paper*.

Buchmueller T.C., Zuvekas S.H. 1998. Drug use, drug abuse, and labor market outcomes. *Health Economics*, 7: 229-245.

Burgess SM, and Propper C. 1998. Early health-related behaviors and their impact on later life changes: evidence from the U.S. *Health Economics*, 7: 381-399.

Chatterji P. 2006. Illicit Drug Use and Educational Attainment. *Health Economics*, 15: 489-511.

Choo E. 2000. Rational Addiction and Rational Cessation: A Dynamic Structural Model of Cigarette Consumption. *Working Paper*.

DeSimone J. 2002. Illegal drug use and employment. *Journal of Labor Economics*, 20: 952-977.

Ellickson P, Bui K., Bell R., McGuigan K. 1998. Does Early Drug Use Increase the Risk of Dropping Out of High School? *Journal of Drug Issues*, 28: 357-380.

Fang H. Silverman D. 2009. Time-inconsistency and welfare program participation: evidence from the NLSY. *International Economic Review*, 50: 1043-1077.

Fang H., Wang Y. 2009. Estimating dynamic discrete choice models with hyperbolic discounting, with an application to mammography decisions. *Working paper*.

Fergusson D.M., Horword L.J. 1997. Early Onset Cannabis Use and Psychosocial Adjustment in Young Adults. *Addiction*, 92: 279-296.

French M., Roebuck C., Alexandre P. 2001. Illicit drug use, employment, and labor force participation. *Southern Economic Journal*, 68. 349-368.

Gill A., Michaels R. 1992. Does drug use lower wages? *Industrial and Labor Relations Review*, 45: 419-434.

Gordon B. Sun B. 2009. A Dynamic Structural Model of Addiction, Promotions, and Permanent Price Cuts. *Marketing Science*, 28: 846-867.

Grossman M, Chaloupka F.J, Sirtalan I. 1998. A Survey of Economic Models of Addictive Behavior. *Journal of Drug Issues*, 28: 631-643.

Gruber J., Koszegi B. Is Addiction Rational. 2001. Theory and Evidence. *Quarterly Journal of Economics*, 116: 1261–304.

Gul F., Pesendorfer W. 2001. Temptation and self-control. *Econometrica*, 69: 1403-1435.

Heckman J, Stixrud J., Urzua S. 2006. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24: 411-482.

Horowitz J. 2001. Should the DEA's STRIDE data be used for economic analyses of markets for illegal drugs? *Journal of the American Statistical Association*, 96: 1254-1262.

Kaestner R. 1991. The effect of illicit drug use on the wages of young adults. *Journal of Labor Economics*, 9: 381-412.

Kaestner R. 1994a. The effect of illicit drug use on the labor supply of young adults. *Journal of Human Resources*, 29: 126-155.

Kaestner R. 1994b. New estimates of the effects of marijuana and cocaine use on wages. *Industrial and labor relations review*, 47: 454-470.

Keane M.P., Wolpin K.I. 1994. The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte Carlo evidence. *Review of Economics and Statistics*, 76: 648-672.

Keane M.P., Wolpin K.I. 1997. Career decisions of young men. *Journal of Political Economy*, 58: 473-522.

Keane M.P., Wolpin K.I. 2001. The effect of parental transfers and borrowing constraints on educational attainment. *International Economic Review*, 42: 1051-1103.

Keane M.P., Sauer R. 2009. Classification error in dynamic discrete choice models: implications for female labor supply behavior. *Econometrica*, 77: 975-991

Keane M.P., Sauer R. 2010. A Computationally Practical Simulation Estimation Algorithm for Dynamic Panel Data Models with Unobserved Endogenous State Variables. *Working Papers 1008, Gutenberg School of Management and Economics*, Johannes Gutenberg-Universität Mainz.

Laibson D. 1997. Golden Eggs and Hyperbolic Discounting. *Quarterly Journal of Economics*, 112: 443–77.

Laibson D., Repetto A., Tobacman J. 2007. Estimating discount functions with consumption choices over the lifecycle. *NBER Working Paper*, 13314.

MacDonald Z. 2004. What price drug use? The contribution of economics to an evidence-based drugs policy. *Journal of Economic Surveys*, 18. 113-151.

MacDonald Z, Pudney S. 2000. Illicit drug use, unemployment, and occupational attainment. *Journal of Health Economics*, 19, 1089-1115.

Machado F., Sinha R. 2007. Smoking cessation: a model of planned vs. actual behavior for time-inconsistent consumers. *Marketing Science*, 26: 834-850.

Manski C.F., Pepper J.V., Petrie C.V. (eds.). 2001. Informing America's policy on illegal drugs: what we don't know keeps hurting us. *Washington DC: National Academy Press.*

McCaffrey D., Pacula R. Han B., Ellickson P. 2008. Marijuana use and high school dropout: the influence of unobservables. *NBER Working Paper*, 14102.

Moffitt R., Gottschalk P. 2008. Trends in the transitory variance of male earnings in the U.S., 1970-2004. *Boston College Working Papers in Economics 697.*

Orphanides A., Zervos D.1995. Rational addiction with learning and regret. *Journal of Political Economy*, 103: 739-758.

Orphanides A., Zervos D.1998. Myopia and addictive behavior. *The Economic Journal*, 108: 75-95.

Pollak R. 1970. Habit formation and dynamic demand functions. *Journal of Political Economy*, 78: 745-763.

Registrer C, Williams D. 1992. Labor market effects of marijuana and cocaine among young men. *Industrial and Labor Relations Review*, 45: 435-448.

Ringel S., Ellickson P., Collins R. 2006. The relationship between high school marijuana use and annual earnings among young adult males. *Contemporary Economic Policy*, 24: 52-63.

Shepard E., Clifton T. 1998. Drug testing and labor productivity: estimates applying a production function model. Le Moyne, France: Le Moyne College Institute of Industrial Relations, 18: 1-30.

Shui H., Ausubel L. 2005. Time inconsistency in the credit card market. *Working paper*.

Stigler R., Becker G. 1977. De gustibus non est disputandum. *American Economic Review*, 67: 76-90.

Thaler R.H., Shefrin H.M. 1981. An economic theory of self-control. *Journal of Political Economy*, 89: 392-406

Van Ours J. 2006. Cannabis, cocaine and jobs. *Journal of Applied Econometrics*, 21: 897-917

Winston G. 1980. Addiction and backsliding: a theory of compulsive consumption. *Journal of Economic Behavior and Organization*, 1: 295-324.

Wolpin K.I., Merlo A. 2008. The Transition from School to Jail: Youth Crime and High School Completion Among Black Males. *Working Paper*.

Poterba and Summers (1986)

Appendix A: Functional Forms

Value Function Approximation

The value function is approximated by a polynomial of second order without the cross-terms. Though the cross-terms increase the precision in the predictions, the gain is low (in terms of adjusted R^2, which is already very high without the cross-terms) compared to the gain in terms of computational time when these terms are not included. More specifically, the approximation contains the following 25 variables: years of completed education (and its squared term), experience (and its squared term), last year school attendance, household type, race, mother's education, non-earned income, education at age 13, a dummy variables to capture if the person consumed soft and hard drugs the previous four periods, 5 cohort-dummy variables, and a constant term.

Terminal Value Function

The variables included to approximate the terminal value function are: average wage the person would get if he decides to work and its squared term; one dummy variable that takes value of one if the person has college education and another one if the person has graduate studies; a dummy variable that takes value of one if the person consumed soft drugs last period and another one that takes value of one if the person consumed drugs from period t-2 to t-4; and same dummy variables for consumption of hard drugs.

Type probability Functions and coefficients that vary by type

Type distribution is conditional on the observed initial conditions (MEd, HH, R, Ed13, NI) and it is specified as a logit. More specifically:

$$\pi_k = \frac{\exp(\phi_{ok} + \phi_{1k} I(M_Ed = 2) + \phi_{2k} I(HH = 1) + \phi_{3k} I(R = 1) + \phi_{4k} I(Ed_13 \geq 7) + \phi_{5k} I(NI > 0))}{1 + \sum_{l=1}^{K} \exp(\phi_{ol} + \phi_{1l} I(M_Ed = 2) + \phi_{2l} I(HH = 1) + \phi_{3l} I(R = 1) + \phi_{4l} I(Ed_13 \geq 7) + \phi_{5l} I(NI > 0))}$$

for k=2,...,K, and $\sum_{k=1}^{K} \pi_k = 1$.

We also allow the coefficients of the utility functions to differ by type.

$$U^{i,k}(c_{it}, l_{it}, a_i^{sdt}, a_i^{hdt}) =$$

$$\frac{c_{it}^{\alpha_{c,k}}}{\alpha_{c,k}} + \frac{l_{it}^{\alpha_{l,k}}}{\alpha_{l,k}} + \{-\alpha_{sd,k} \exp[(\sum_{p=t-1}^{T_0} (\beta_{sd,k})^p a_{ip}^{sd} - a_{it}^{sd})/\alpha_{sd,k}] + \varepsilon_{it}^{sd} I(a_{it}^{sd} > 0)\}$$

$$+\{-\alpha_{hd,k} \exp[(\sum_{p=t-1}^{T_0} (\beta_{hd,k})^p a_{ip}^{hd} - a_{it}^{hd})/\alpha_{hd,k}] + \varepsilon_{it}^{hd} I(a_{it}^{hd} > 0)\}$$

with k=1,...,K.

Thus, different types can differ in their aversion to risk, or in their intertemporal elasticity of substitution for leisure, or in their preferences for drugs. This last difference tries to capture the fact that some people can like drugs more than others for unobserved reasons.

We also allow the constant term of the wage equation to differ by type to capture differences in unobserved permanent skill endowments that existed at age 13.

Finally, we allow the mean of the stochastic draws for soft and hard drugs to differ by type. More specifically,

$$\mu^{k}_{\varepsilon^{sd}_{it}} = \rho_{ok} + \rho_1 MEd + \rho_2 HH + \rho_3 R + \rho_4 Ed13$$

with k=2,...,k.

This would capture that there are types of individuals who are more prone to get draws of the errors that make them more or less likely to consume drugs than other types. For example, if some individuals tend to be surrounded by peers that act as a "bad influence", that is, that consume drugs (which is an unobserved factor in the model) and this make them more likely to receive negative shock, the differences in ρ_{ok} would capture this permanent unobserved difference.

Appendix B: Tables and Graphs

Table B.1. Summary Statistics

Variable	Obs.	Mean	Std. Dev	Min.	Max.
Highest Grade Completed	41323	10.40	2.44	5	19
School Attendance in t	41016	0.64	0.48	0	1
School Attendance in t-1	42751	0.69	0.46	0	1
Soft Drug Cons. in t	40010	0.22	0.42	0	1
Hard Drug Cons. In t	40439	0.06	0.23	0	1
Non-Earned Income (age 13-15)*	5953	0.69	0.81	0	2
Non-Earned Income (age 16-25)*	28039	1.36	1.09	0	3
Household Type	4198	0.52	0.50	0	1
Race	4203	0.50	0.50	0	1
Education at Age 13**	4197	0.69	0.46	0	1
Mother's Education	3970	1.41	0.49	1	2
Soft Drug Cons. in t-1	40707	0.20	0.40	0	1
Soft Drug Cons. in t-1	41327	0.17	0.38	0	1
Soft Drug Cons. in t-2	41751	0.15	0.36	0	1
Soft Drug Cons. in t-2	42110	0.12	0.33	0	1
Soft Drug Cons. in t-3	41022	0.05	0.22	0	1
Soft Drug Cons. in t-3	41715	0.05	0.21	0	1
Soft Drug Cons. in t-4	42333	0.04	0.20	0	1
Soft Drug Cons. in t-4	42759	0.03	0.18	0	1
Work	28611	0.67	0.47	0	1
Hourly Wage	16730	4.99	2.46	1.016	25
Job Experience	27960	2.11	2.17	0	9
Wealth in t	9130	10,214	26,444	-517,632	356,506
Wealth in t-1	13043	6,917	21,372	-107,220	356,506

*Non-earned Income is expressed in categories that range from 0 to 2 for ages 13 to 15, and 0 to 3 for individuals older than 15.

**Education at age 13 takes a value of 1 if the individual is in grade 6, 2 if in grade 7, 3 if in grade 8, and 0 if in grade 5. However, for this table it was coded as 1 if the individual is in grade 7 or 8, and 0 otherwise.

Table B.2. Age of Initiation into Soft and Hard Drugs

	Soft Drugs			Hard Drugs		
Age	Obs	%	Acum. %	Obs	%	Acum. %
1-6	19	0.45	0.45	11	0.26	0.26
7-10	47	1.12	1.57	14	0.33	0.59
11	54	1.28	2.86	15	0.36	0.95
12	136	3.24	6.09	35	0.83	1.78
13	235	5.59	11.68	65	1.45	3.24
14	339	8.07	19.75	84	2.00	5.23
15	365	8.68	28.43	127	3.02	8.26
16	337	8.02	36.45	155	3.69	11.94
17	334	7.95	44.40	143	3.40	15.35
18	270	6.42	50.82	124	2.95	18.30
19	179	4.26	55.08	94	2.24	20.53
20	115	2.74	57.82	77	1.83	22.36
21	79	1.88	59.70	66	1.57	23.94
22	60	1.43	61.12	34	0.81	24.74
23	32	0.76	61.88	21	0.50	25.24
24	19	0.45	62.34	4	0.10	25.34
25	11	0.26	62.60	2	0.05	25.39

Table B.3. Distribution of Drug Consumption by Age

	No Drugs	Only Soft Drugs	Only Hard Drugs	Soft & Hard Drugs
Age	%	%	%	%
13	92.04	5.58	1.44	0.94
14	86.80	9.71	2.20	1.29
15	81.73	13.36	2.28	2.63
16	75.61	17.43	2.63	4.32
17	70.08	22.18	2.34	5.40
18	68.56	24.39	1.73	5.32
19	69.19	24.20	1.44	5.17
20	70.60	22.95	1.51	4.93
21	72.00	21.56	2.07	4.36
22	74.28	19.92	1.89	3.91
23	76.27	18.64	1.18	3.91
24	77.32	17.42	1.44	3.81
25	78.39	17.29	1.73	2.59

Table B.4. Years of Completed Education

Age	Non-Drug Users	Marijuana Users	Hard Drugs Users
13	6.74	6.67	6.68
14	7.75	7.73	7.71
15	8.73	8.67	8.64
16	9.68	9.54	9.54
17	10.59	10.36	10.32
18	11.47	11.06	10.98
19	12.07	11.49	11.39
20	12.49	11.84	11.74
21	12.85	12.10	12.02
22	13.09	12.36	12.21
23	13.34	12.46	12.36
24	13.41	12.66	12.49

Table B.5. School Attendance

Age	Non-Drug Users	Marijuana Users	Hard Drugs Users
13	99.90	100.00	100.00
14	99.75	98.74	98.89
15	98.79	95.60	94.37
16	96.08	86.63	85.50
17	90.48	71.92	69.84
18	64.78	46.25	42.29
19	51.48	34.77	31.56
20	46.09	30.57	27.95
21	42.61	28.37	25.00
22	32.71	20.58	18.16
23	24.93	15.79	15.82
24	16.78	16.77	13.50

Table B.6 Utility function parameters

α_c	α_l	α_{sd}	β_{sd}	α_{hd}	β_{hd}
0.564	0.765	4.163	0.113	8.015	0.633
(0.000021)	(0.000028)	(0.000585)	(0.000015)	(0.004674)	(0.000368)

Table B.7. Wage Equation and School Related Parameters

β_1	β_2*	β_3*	β_4*	β_5*
0.164537	0.1598	0.6164	1.5924	-0.0127
(0.000069)	(0.01)	(0.03)	(0.003)	(0.01)
β_6	β_7*	β_8*	β_9*	β_{10}*
0.099825	-0.5147	-1.1996	-0.0746	-1.4018
(0.000005)	(0.001)	(0.003)	(0.002)	(0.001)
β_{11}	β_{12}*	β_{13}*	β_{14}*	β_{15}*
0.067324	0.0166	2.994	-0.0231	-0.7853
(0.000003)	(0.001)	(0.02)	(0.001)	(0.001)
β_{16}**	cs_u	cs_g	γ_1***	γ_2
-1.02656	6080.8	11646.7	1040	315.06
(0.001)				(0.101611)

*coefficient and standard errors multiplied by 10,000
**coefficient and standard errors mutiplied by 1,000
***this coefficient is a constant in the model (not estimated)

Table B.8. Variance-Covariance Matrix

σ_s	σ_w	σ_{sd}	σ_{hd}	$\rho_{s,w}$
5852.240	0.189	66.418	14.856	-0.039
(0.553437)	(0.01324)	(0.009893)	(0.000553)	(0.000022)
$\rho_{s,sd}$	$\rho_{s,hd}$	$\rho_{w,sd}$	$\rho_{w,hd}$	$\rho_{sd,hd}$
-0.042	0.124	-0.086	0.020	0.978
(0.000172)	(0.000511)	(0.000356)	(0.000842)	(0.000013)

Table B.9. Initial Condition Parameters

Marijuana		
Race	Household Type	6 years of Educ. Age 13
0.007219	-0.063567	0.133016
(0.021950)	(0.021816)	(0.039816)
Marijuana		
7 years of Educ. Age 13	8 years of Educ. Age 13	Mother's Educ.
-0.004936	-0.078018	-0.083032
(0.035921)	(0.012955)	(0.013653)
Hard Drugs		
Race	Household Type	6 years of Educ. Age 13
0.000625	0.000625	0.977112
(0.025135)	(0.025243)	(0.069411)
Hard Drugs		
7 years of Educ. Age 13	8 years of Educ. Age 13	Mother's Educ.
-0.166567	0.044917	0.02366
(0.036721)	(0.317777)	(0.000884)

Table B.10. Classification Error Rate Parameters and Measurement Error in Hourly Wages

School	Work	Marihuana	Other Drugs	Mother's Educ.
0.372135	0.71321	0.756642	0.392411	0.118903
(0.006342)	(0.005092)	(0.005158)	(0.011808)	(0.044205)
Race	Household Type	Non-Earned Inc.	Educ. At Age 13	Hourly Wage
0.150396	0.187353	0.094932	0.153983	0.181829
(0.040918)	(0.040937)	(0.074753)	(0.014908)	(0.002103)

Table B.11. Probability of Repetition

η_0	η_1	η_2	η_3	η_4
-7.468	0.293	-0.645	-0.162	0.479
(0.369)	(0.063)	(0.109)	(0.039)	(0.035)
η_5	η_6	η_7	η_8	
0.220	0.268	0.420	0.118	
(0.076)	(0.118)	(0.070)	(0.098)	
η_9	η_{10}	η_{11}	η_{12}	
-0.713	-0.473	-0.432	-0.0006	
(0.060)	(0.060)	(0.059)	(0.057)	

Table B.13. Returns to School and Work Experience

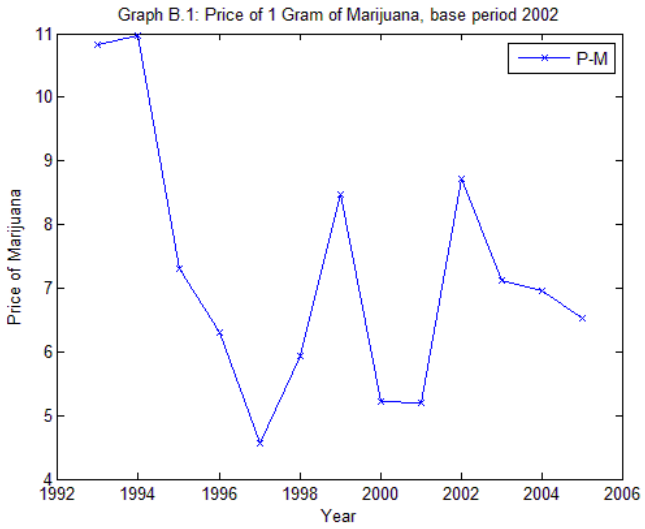
	Non-Drug Consumer	Marijuana Consumption			Hard Drugs Consumption	
		12 days per year	120 days per year	300 days per year	8 times per year	52 times per year
Returns to School	9.98	9.92	9.39	8.49	9.45	6.51
Returns to Work	6.73	6.72	6.64	6.51	6.72	6.64

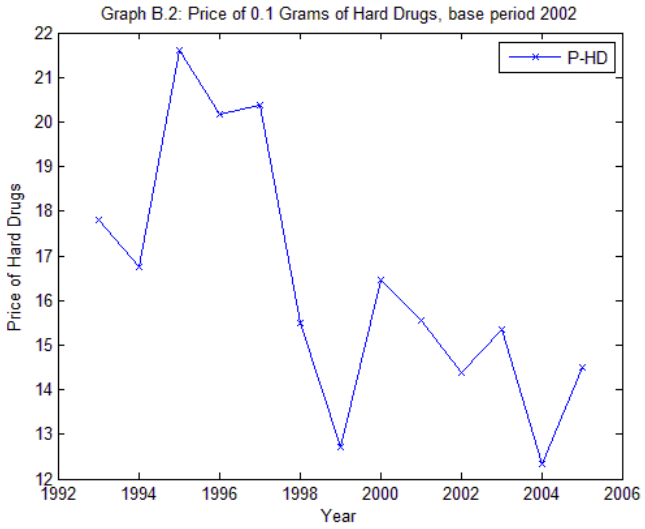
Table B.14. Hourly Wages by Levels of Drug Consumption

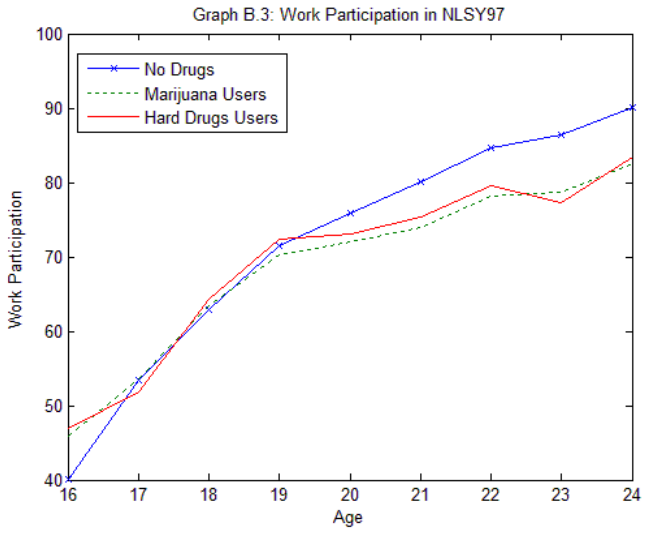
	Non-Drug Consumer	Marijuana Consumption			Hard Drugs Consumption	
		12 days per year	120 days per year	300 days per year	8 times per year	52 times per year
Complete College*	7.50	7.43	6.86	5.99	6.89	4.34
Complete Highschool*	5.03	5.00	4.71	4.27	4.72	3.34
Complete College**	12.22	12.10	11.08	9.56	11.22	7.00
Complete Highschool**	8.20	8.14	7.61	6.80	7.69	5.40

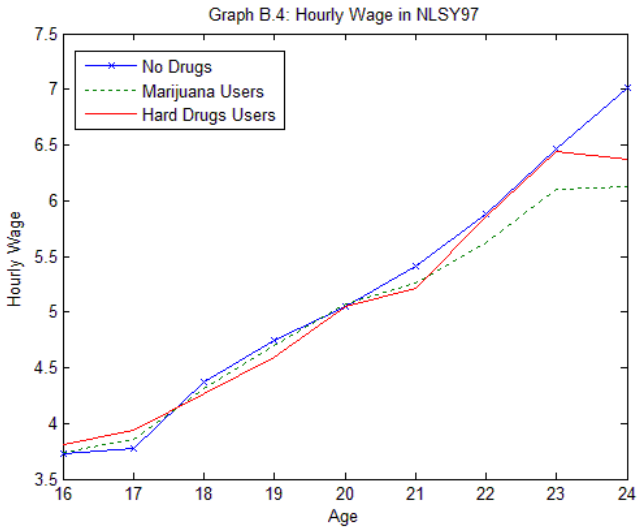
*Considering work experience equal to 4 years

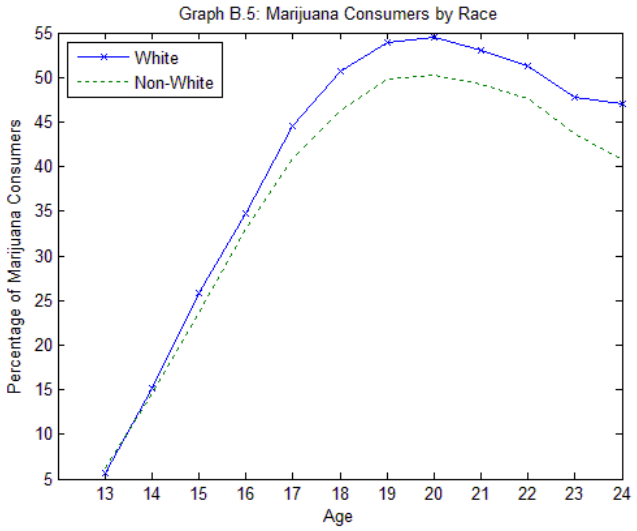
**Considering work experience equal to 14 years

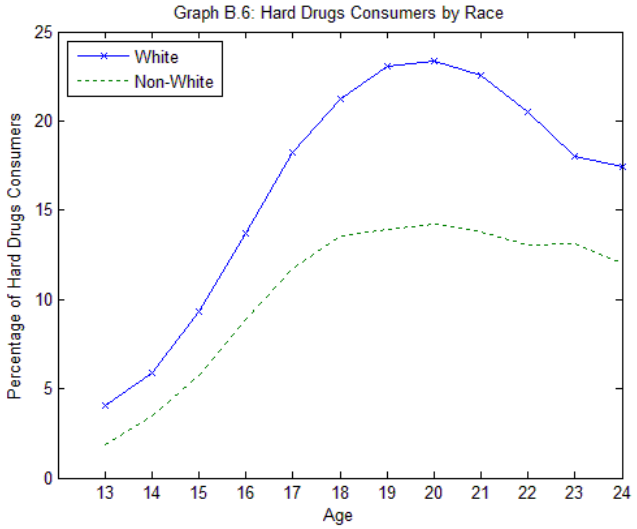

Graph B.1: Price of 1 Gram of Marijuana, base period 2002
P-M
Price of Marijuana
Year
11
10
9
8
7
6
5
4
1992
1994
1996
1998
2000
2002
2004
2006

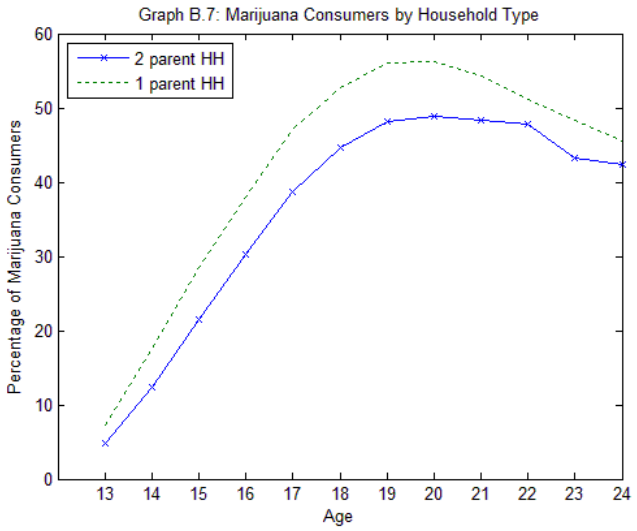

Graph B.2: Price of 0.1 Grams of Hard Drugs, base period 2002
P-HD
Price of Hard Drugs
Year
22
21
20
19
18
17
16
15
14
13
12
1992
1994
1996
1998
2000
2002
2004
2006

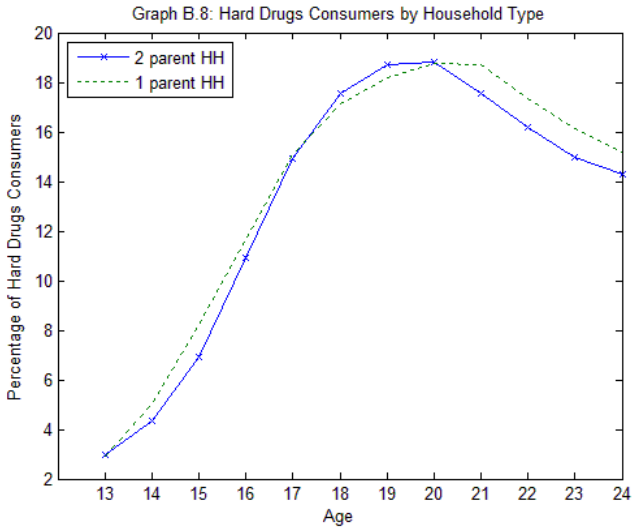

Graph B.3: Work Participation in NLSY97
No Drugs
Marijuana Users
Hard Drugs Users
Work Participation
Age
100
90
80
70
60
50
40
16
17
18
19
20
21
22
23
24

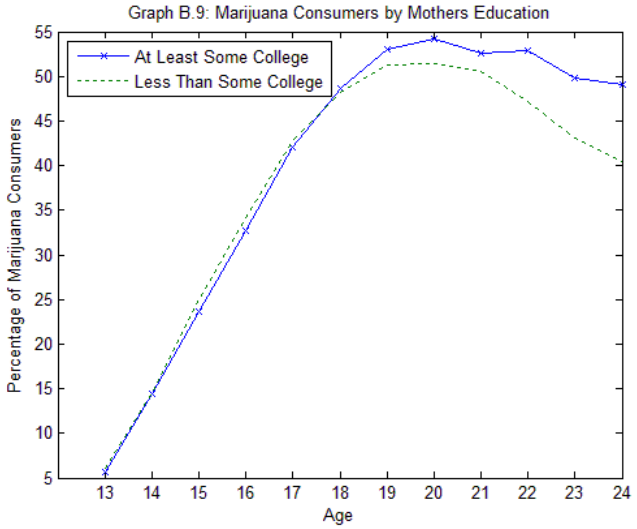

Graph B.4: Hourly Wage in NLSY97
No Drugs
Marijuana Users
Hard Drugs Users
Hourly Wage
Age
7.5
7
6.5
6
5.5
5
4.5
4
3.5
16
17
18
19
20
21
22
23
24


Graph B.5: Marijuana Consumers by Race
White
Non-White
Percentage of Marijuana Consumers
Age
55
50
45
40
35
30
25
20
15
10
5
13
14
15
16
17
18
19
20
21
22
23
24

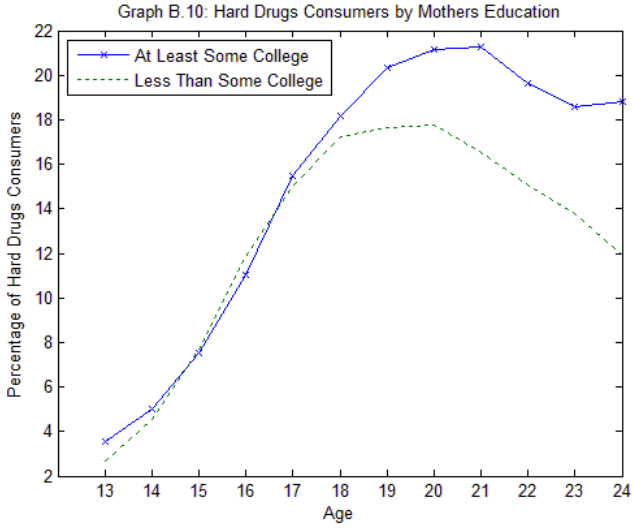

Graph B.6: Hard Drugs Consumers by Race
White
Non-White
Percentage of Hard Drugs Consumers
Age
25
20
15
10
5
0
13
14
15
16
17
18
19
20
21
22
23
24


Graph B.7: Marijuana Consumers by Household Type
2 parent HH
1 parent HH
Percentage of Marijuana Consumers
Age
0
10
20
30
40
50
60
13
14
15
16
17
18
19
20
21
22
23
24

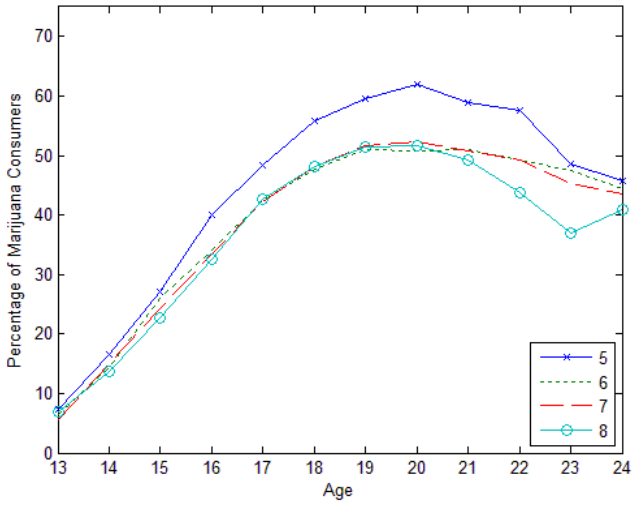

Graph B.8: Hard Drugs Consumers by Household Type
2 parent HH
1 parent HH
Percentage of Hard Drugs Consumers
Age
2
4
6
8
10
12
14
16
18
20
13
14
15
16
17
18
19
20
21
22
23
24


Graph B.9: Marijuana Consumers by Mothers Education
At Least Some College
Less Than Some College
Percentage of Marijuana Consumers
Age
5
10
15
20
25
30
35
40
45
50
55
13
14
15
16
17
18
19
20
21
22
23
24

Graph B.10: Hard Drugs Consumers by Mothers Education



Graph B.11:Marijuana Consumers by Initial Level of Schooling



Graph B.12: Hard Drugs Consumers by Initial Level of Schooling

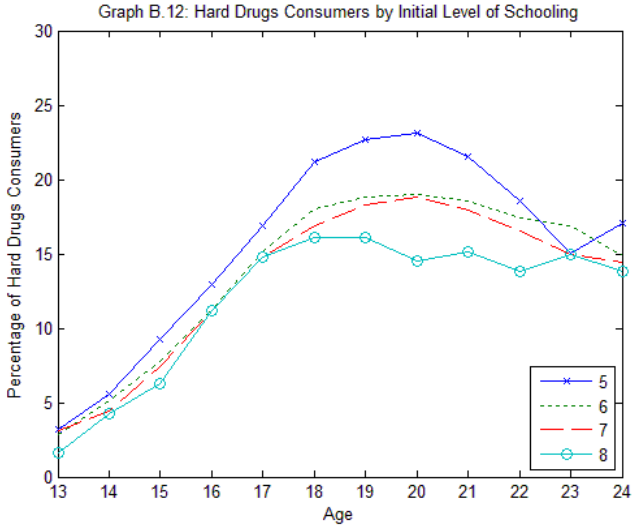


Fig. B.13: School Attendance in NLSY97

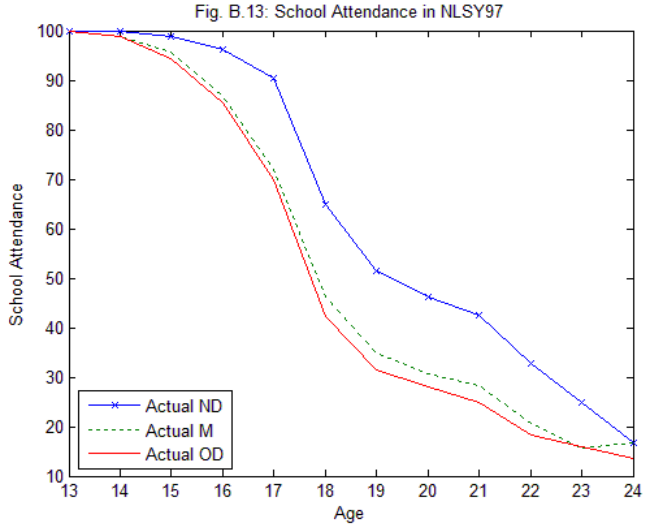


Fig. B.14: School Attendance in Predicted Data

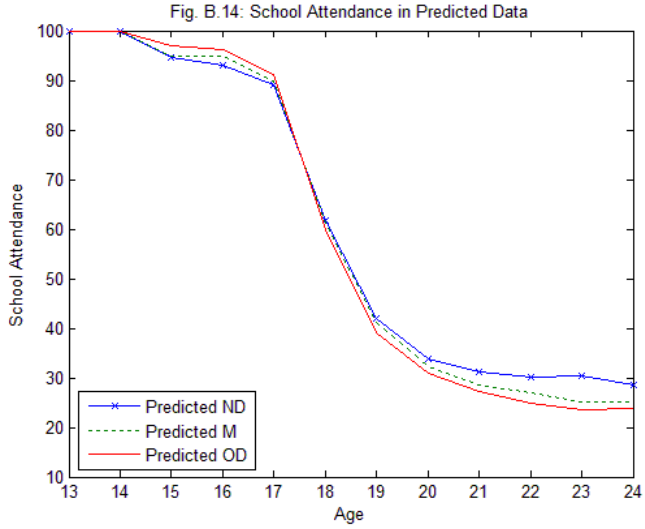


Fig B.15: Work Participation in Predicted Data

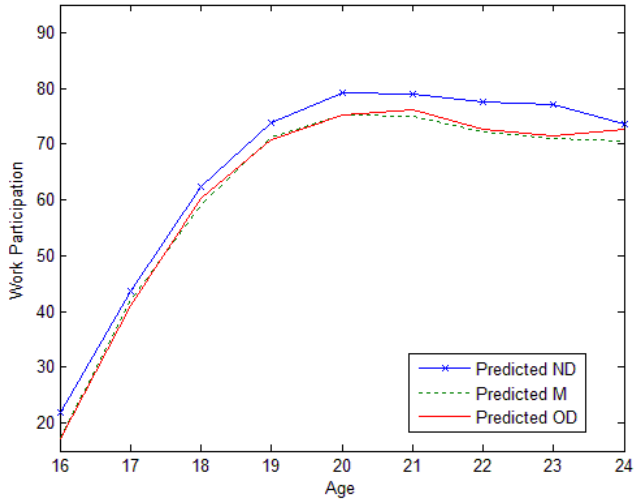


Fig B.16: Hourly Wage in Predicted Data

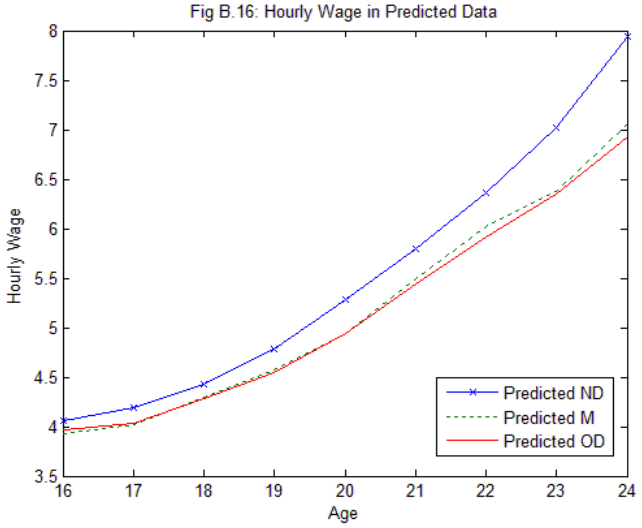


Fig. B.17: Percentage of Current Marijuana Consumption in Actual and Predicted Data

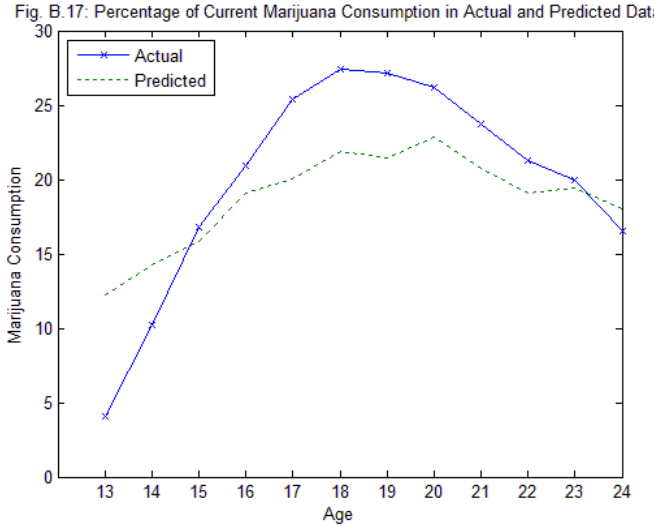
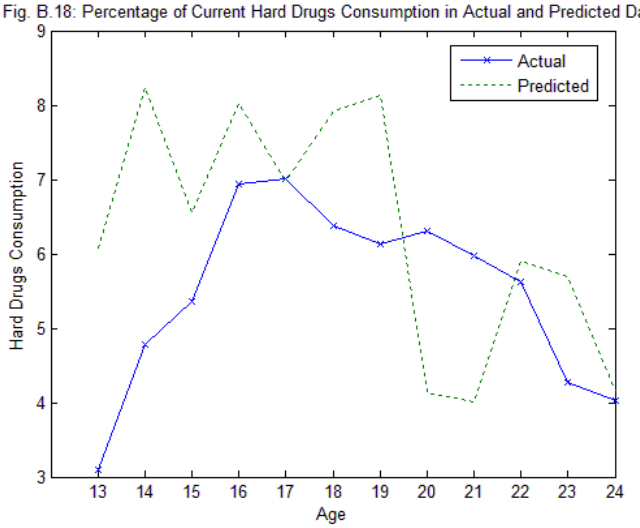
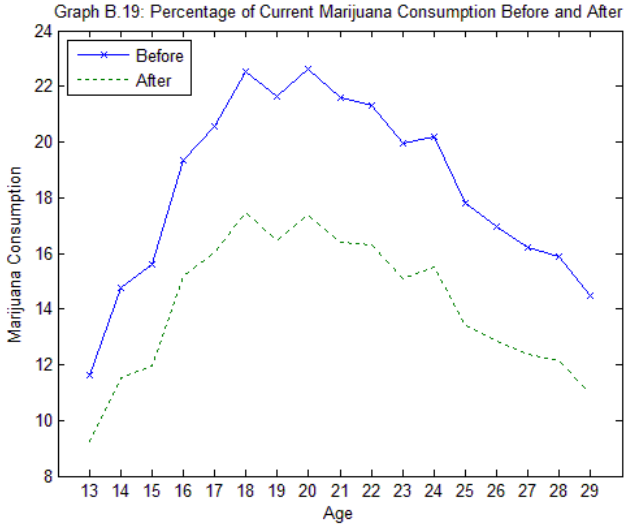


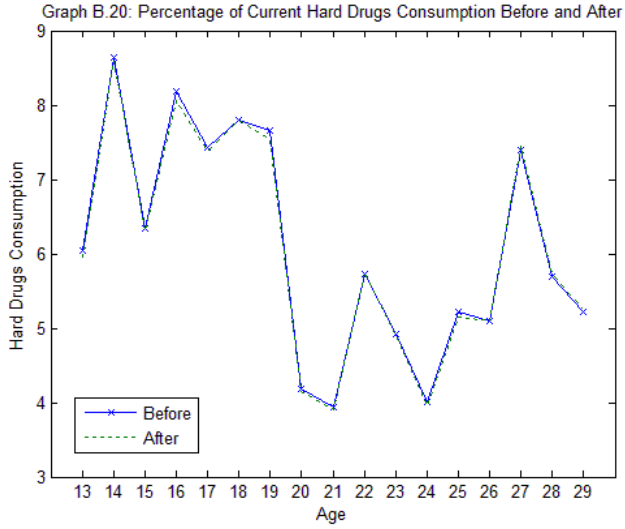
Fig. B.18: Percentage of Current Hard Drugs Consumption in Actual and Predicted Data



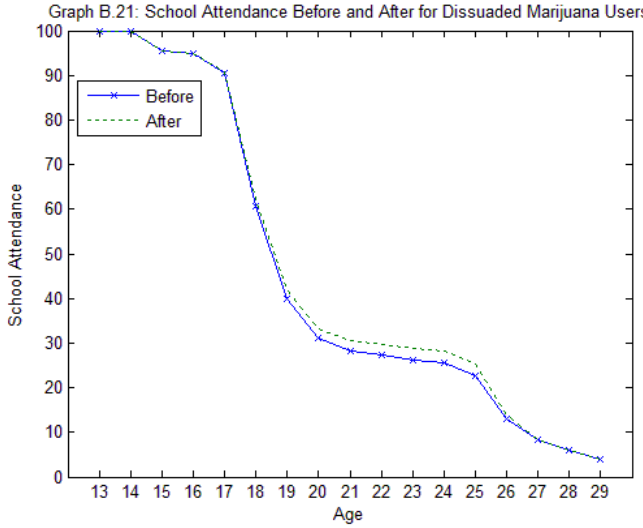
Graph B.19: Percentage of Current Marijuana Consumption Before and After



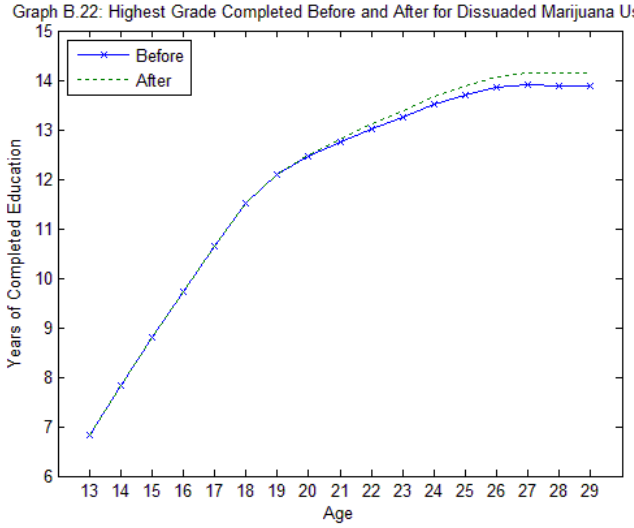
Graph B.20: Percentage of Current Hard Drugs Consumption Before and After



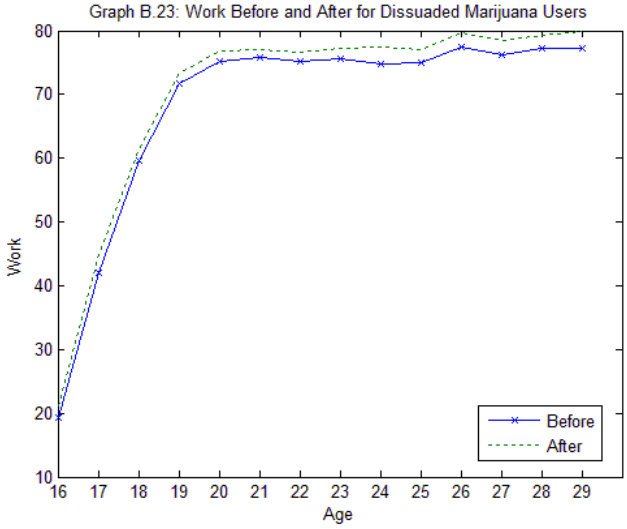
Graph B.21: School Attendance Before and After for Dissuaded Marijuana Users



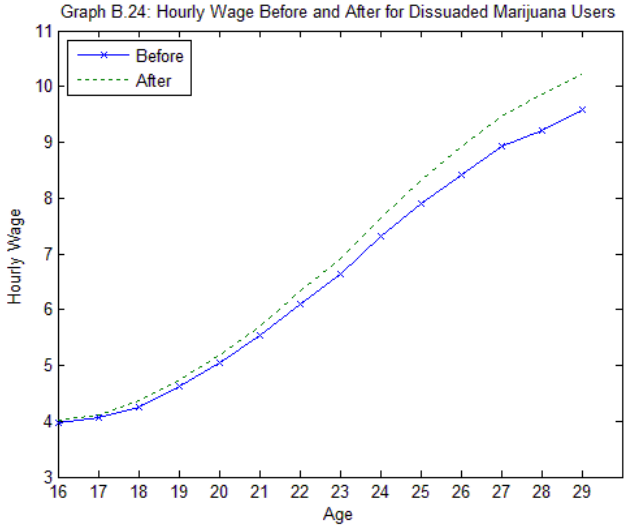
Graph B.22: Highest Grade Completed Before and After for Dissuaded Marijuana Users



Graph B.23: Work Before and After for Dissuaded Marijuana Users



Graph B.24: Hourly Wage Before and After for Dissuaded Marijuana Users



7 Appendix C: More on Definition of Choices and Variables

Prices of Drugs

Constructing prices for hard drugs is complex since the NLSY97 does not allow frequency of consumption to be disaggregated for different types of hard drugs. The STRIDE data does not have information for other drugs like hallucinogens, inhalants, etc, therefore we only considered the prices of crack, cocaine and heroin to construct a series for the price of hard drugs. This omission to construct this series is not problematic as long as consumption and prices of these additional drugs follow a pattern similar to the ones included in the analyses. We use the prices for one gram of each drug reported by Arkes et al. (2008), which are calculated using data for Washington DC.[32] Given that one gram of these drugs is too much for one time consumption, we adjust the prices assuming that when a person reports consuming these drugs for a day, he is consuming 0.1 grams of cocaine and crack (which is approximately the amount of cocaine in two lines) and 0.042 grams of heroin. This amount of heroin comes from a study by Agar et al. (1998), who follow heroin addicts in three cities: San Francisco, CA, Baltimore, MD, and Newark, NJ, and report drug consumption per day for low, medium and high habit heroin users. They find that low habit heroin users consume 14mg of the drug per day in San Francisco, 62mg in Baltimore, and 51mg in Newark (which averages to 0.042 grams). We choose average consumption of low habit users because this conforms to reported drug consumption in the NLSY97. To aggregate the three prices, we weight the price of each drug by how much that drug represents in the total consumption of hard drugs. These weights are taken from the National Survey on Drug Use and Health 2006 and are disaggregated by age group.[33] Finally, it is assumed that all the people born in the same year face the same price in a given year.[34] All prices are in year 1983 dollars.

Non-earned Income

Respondents were asked a set of questions about non-earned income starting from round one of the survey. These questions are about income from illegal sources (including property, crime, and illegal drugs), allowance from the family, money other than allowance received from different sources that the individual is not expected to repay, and financial aid received in grants, tuition or fee waivers or reductions, and fellowships or scholarships while attending college. This variable was discretized in 3 categories for ages 13 to 15 and 4 categories for older ages. The three categories for ages 13 to 15 are given by: (a) \$0-\$5, (b) \$5-\$200, (c) \$200-\$50,000. For older ages, the 4 categories are: (a) \$0-\$5, (b) \$5-\$200, (c) \$200-\$1,000, (d) \$1,000-\$50,000. Values above \$50,000 were coded as missing. The specific value for each

[32] The authors note the importance of discounts per quantity in the drugs market. To consider this, they calculate the price of a gram of crack, of cocaine, and of heroin for different quantities. In this analysis I consider the prices corresponding to the smallest amounts (0.3 grams for crack, 0.75 grams for cocaine, and 0.4 grams for heroin), which are small enough to be considered as quantities a final consumer buys. The price of crack corresponds to crack that is 75 percent pure, while the prices for cocaine and heroin correspond to 100 percent purity

[33] This information can be found on the official site for the database: http://www.oas.samhsa.gov/nhsda.htm, in the Tables section. For ages above 25 years old, I use the percentages for the age group 26-34.

[34] Since there is no information for prices beyond 2005, I use 2005 prices for subsequent years.

category is given by the median value for each age. To calculate the transition probabilities individuals use to form expectations about the future while solving the dynamic problem, we aggregated the data in three groups: (i) those aged 13 and 14; (ii) those aged 15; and (iii) those aged 16 and older. This decision was taken because the distributions per category for each of these age groups look very similar (but different across groups).

Other Variables

Race is a dummy variable which equals one if the person is white and zero otherwise. Household type at age 12 is a dichotomous variable that equals one if the person lived with both biological parents at age 12 or with at least with one biological parent in a two-parent household, and zero otherwise. Mother's education is a dummy variable that equals one if the mother has at least college education and zero otherwise.