High Today vs Lows Tomorrow:
Substance Use, Education, and Employment Choices of Young Men

Catherine Elizabeth Alford
Jacksonville, Florida

M.A. Economics, University of Virginia, 2010
B.A. Economics, University of Florida, 2008
B.S. Mathematics, University of Florida 2008

A Dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of Doctor of Philosphy

Department of Economics

University of Virginia
May, 2015

Abstract

In this paper, I develop and estimate a dynamic structural model of education, employment, and substance use decisions of young men in order to determine the causal effects of substance use on educational attainment and career paths. Heavy substance use is correlated with lower school attainment and labor market outcomes; however, it is unclear if heavy substance use *causes* these worse outcomes. One concern is that those who are more likely to use marijuana or alcohol frequently are those for whom the labor market would offer lower wages regardless of their substance use. I utilize variation in the prices of substances, the price of college, local law enforcement characteristics, and unemployment rates to help identify the channels through which current substance use and school decisions affect future substance use, employment decisions, and wages. Current research generally treats the substance use decision as a binary choice, making it difficult to distinguish the effects of moderate versus heavy use. I allow individuals to make choices about their levels of alcohol, cigarette, and marijuana use in order to capture the full relationship between substance use and outcomes.

I estimate my model using restricted-access data from the 1997 National Longitudinal Survey of Youth and Bayesian Markov Chain Monte Carlo (MCMC) methods. My model includes extensive time-persistent unobserved heterogeneity, which helps me to estimate the causal effects of substance use on outcomes. I use a modified version of the estimator proposed in Imai, Jain, and Ching (2009) in order to make estimation feasible by easing the computational burden of evaluating my likelihood function and value functions.

I find that heterogeneity in preferences plays an important role in an individual's choice to use cigarettes and marijuana. That is, without heterogeneity, the proportion of individuals using cigarettes and marijuana would be much less. I find that cigarette and alcohol

use have causal effects that decrease the wages of white and Hispanic males, but have no statistically significant effects on the wages of black males. I also find that past and present alcohol and cigarette use affect an individual's choice to use marijuana, with alcohol having a larger effect. Additionally, I find that marijuana use leads individuals into heavy cigarette use, supporting the reverse gateway theory.

Lastly, I find that white males have a lower probability of arrest than black and Hispanic males, conditional on age, previous arrests, and substance use. In addition, the use of heavy marijuana increases the probability of arrest more for black males than it does for white or Hispanic males. White males are also more likely to graduate from high school and to be working full-time at the age of 24, even though their substance use is comparable to that of black and Hispanic males. This suggests that arrests may be contributing to the education gap between white and minority males. I run two policy simulations to see if the outcomes of minority males can be improved by decreasing the probability of arrest. In the first, I set the coefficients in the probability of arrest equations of Hispanic and black males equal to those in the equation of white males. I find that high school graduation rates increase by 3.8 percent and 6.7 percent for Hispanic and black males, respectively. I also find that the proportion of black males using heavy amounts of substances decreases substantially. In the second policy simulation, I consider the effects of decreasing the marginal effect of marijuana use on the probability of arrest of black males, through, for example, legalizing marijuana. This policy change has no effect on high school graduation rates, but I similarly find that decreasing the probability of arrest decreases the proportion of black males using heavy amounts of cigarettes, alcohol, or marijuana.

Acknowledgements

I have received support and encouragement from a great number of individuals. I am especially grateful for the time and support Steven Stern, Leora Friedberg, and Sarah Turner have given me as I developed this paper from an idea to a completed study. Their guidance and encouragement has made this a thoughtful and rewarding journey.

I would like to thank the University of Virginia's Bankard fund for financial support while writing this dissertation. I would also like to thank Andrew Ching, Victor Aguirregabiria, Aloysius Siow, and the participants of the SWEAT workshop at the University of Toronto for their helpful suggestions. My work has also benefited from valuable comments from Jennifer Doleac, Ed Olsen, John Pepper, William Johnson, Amalia Miller, and Matthew Harris.

I will forever be indebted to my friends and classmates who have supported me during this process. My work has benefited from countless conversations with many classmates including Ignacio Martinez, Charlie Murry, Patten Mahler, Kelli Bird, Dusan Curcic, and Chris Clapp. I am also grateful for Susan Clapp, Stephanie Demperio, and Becky Tippett for always being there to encourage me.

Finally, and most importantly, I would like to thank my mom and dad for reminding me that all I can do is my best and that I need to be proud of myself no matter the outcome (and maybe a little extra proud when it actually works out); my sister for always picking up the phone when I need to talk; and my husband for his unwavering love and support. I would not have been able to accomplish this without them.

This dissertation is dedicated to Nate.

Contents

1 Introduction

Substance use is prevalent among American youth. In 2012, 6.8 percent of 12 to 17 year old males reported smoking cigarettes in the past month, 12.6 percent reported drinking, and 7.5 percent reported using marijuana.[1] Heavy substance use is correlated with lower school attainment and labor market outcomes; however, it is unclear if heavy substance use *causes* these worse outcomes. One concern is that those who are more likely to use, say, marijuana frequently are those for whom the labor market would offer lower wages regardless of marijuana use. It may also be the case that poor expected labor market outcomes make substance use less costly.

In this paper, I develop a model that captures these channels relating substance use, educational attainment, and career paths of young men. I develop a dynamic structural model where individuals make decisions about schooling and work, as well as how much alcohol, cigarettes, and marijuana to consume. My model allows substance use to affect career paths through its effects on educational attainment, criminal record, and wage offers. I allow substance use to influence educational attainment by affecting whether an individual enrolls in school as well as whether an individual advances to the next grade if he enrolls. Individuals make decisions in order to maximize their discounted lifetime utility, accounting for the effects of substance use on future outcomes. Individuals make different choices due to both observed characteristics, such as prices, and unobserved characteristics, such as their individual predilection for substance use.

A large literature, both reduced-form and structural, examines how substance use affects educational attainment and wages. Typically, research on this subject focuses on only one substance and one outcome. However, youth who engage in one form of risky behav-

[1] 2012 National Survey on Drug Use and Health.

ior are likely to participate in other forms as well. For example, consider the finding that youth who smoke cigarettes have lower educational attainment than those who do not. If individuals who smoke cigarettes do so only when they consume alcohol, then perhaps it is the alcohol that causes the lower educational attainment and not the cigarettes. Moreover, the current literature generally limits the substance use decision to a binary choice. Yet, it is commonly believed that heavy substance use adversely affects education and career outcomes more than moderate use does. I allow individuals to make choices about their level of substance use to capture the full relationship between substance use levels and outcomes.

I estimate my model using Bayesian Markov Chain Monte Carlo (MCMC) methods. My model includes extensive unobserved heterogeneity and continuous state variables, which make evaluating the likelihood function computationally difficult. I use a modified version of the estimator proposed in Imai, Jain, and Ching (2009) (IJC) in order to make estimation feasible by easing the computational burden of evaluating my likelihood and value functions. I use restricted-access data from the 1997 National Longitudinal Survey of Youth (NLSY97) to estimate my model. The NLSY97 collects information about labor market behavior and educational experiences, as well as information on substance use. Variation in the prices of substances, college prices, law enforcement characteristics, and unemployment rates help identify the channels through which current substance use and schooling decisions affect future substance use, employment decisions, and wages. Most papers in the literature use marijuana prices from the Drug Enforcement Administration's (DEA's) System to Retrieve Information from Drug Evidence. The DEA's focus is on harder drugs like heroin, so the number of marijuana observations are small. I use marijuana price data assembled from High Times magazine, which has much better geographic coverage. Details of how I assemble these prices can be found in Alford (2013).

I find that the effects of substance use on outcomes vary across race, the type of substance used, and the amount used. I find that exogenously consuming heavy amounts of cigarettes decreases the utility of enrolling in school for all males; exogenously consuming heavy amounts of alcohol increases the utility of enrolling for Hispanic males, but has no statistically significant effect on white or black males; and exogenously consuming heavy amounts of marijuana decreases the utility of enrolling for Hispanic males, but increases it for white and black males. Cigarette and alcohol use do not affect the probability that an individual passes a grade. On the other hand, heavy marijuana use decreases the probability that a black male will pass a grade. For all races, moderate alcohol use increases the utility of working, perhaps due to the social aspects of consuming a moderate amount of alcohol. I find that cigarette and alcohol use decrease the wages of white and Hispanic males, but have no statistically significant effect on the wages of black males. Additionally, I find that marijuana use has no significant effect on wages. I find that past and present alcohol and cigarette use affect an individual's choice to use marijuana, with alcohol having a larger effect. Lastly, I find evidence of the reverse gateway theory. That is, I find that marijuana use leads individuals into heavy cigarette use.

I find that unobserved heterogeneity in preferences and skills also vary across race. For example, for white and Hispanic males, having a higher than average preference for heavy marijuana use is correlated with higher preferences for working, higher chances of getting arrested, lower chances of passing a grade, and higher earnings. For black males, having higher preferences for heavy marijuana use is correlated with having lower than average preferences for working part-time, higher preferences for enrolling in school, lower probability of passing, and higher wages. I run a counterfactual simulation that eliminates heterogeneity in preferences and ability to demonstrate the importance of unobserved preferences on the choice to use substances. I find that a large part of an individual's choice to use cigarettes and marijuana is due to heterogeneity in preferences. That is, without heterogeneity, the proportion of individuals using cigarettes and marijuana would be much less. Heterogeneity in preferences also plays an important role in the choices of white and black men to use heavy amounts of alcohol.

Lastly, I find that white males have a lower probability of arrest than black and Hispanic males, conditional on age, previous arrests, and substance use. I also find that the marginal effect of substance use on arrest probabilities differs across race. Using a heavy amount of marijuana increases the probability of arrest more for black males than it does for white or Hispanic males; heavy alcohol use disproportionately increases the arrest probability of Hispanic males. In addition to facing lower arrest rates, white males are also more likely to graduate from high school and to be working full-time at the age of 24 even though their substance use is comparable to the use of black and Hispanic males. This suggests that arrests may be contributing to the education gap between white and minority males. I run two policy simulations aimed at improving outcomes of minority males by changing the probability of arrest. In the first, I set the coefficients in the probability of arrest equations of Hispanic and black males equal to those in the equation of white males. I find that high school graduation rates increase by 3.8 percent and 6.7 percent for Hispanic and black males, respectively. I also find that the proportion of black males using heavy amounts of substances decreases. In the second policy simulation, I consider the effects of decreasing the marginal effect of marijuana use on the probability of arrest of black males, through, for example, legalizing marijuana. To do this, I set the coefficients on moderate and heavy marijuana in the probability of arrest equation equal to the coefficients on moderate and heavy alcohol. I do not find any effect on high school graduation rates, but I find that decreasing the marginal effect of marijuana use on the probability of arrest of black males decreases the proportion using heavy amounts of cigarettes, alcohol, or marijuana.

2 Literature Review

My paper builds on previous work in two areas of the literature. First, my paper is motivated by the substance use literature, which includes papers that analyze the determinants of substance use and papers that analyze the effects of use on labor market and educational outcomes. Second, my model expands on those developed in the human capital investment literature, particularly involving dynamic labor supply, by modeling leisure decisions of individuals.

2.1 Substance Use Literature

Rational Addiction Literature: Most empirical studies of the rational addiction theory follow Becker and Murphy (1988). They develop a theory in which individuals make decisions about consuming goods that they know to be addictive by weighing how their choices today will affect their future utility. Chaloupka (1991) and Becker, Grossman, and Murphy (1994) were two of the first papers to empirically test this theory and find evidence supporting it. Labeaga (1999) improves on this by estimating a rational addiction model for tobacco that incorporates serial correlation and unobserved heterogeneity. The empirical results reject the myopic model and give support to the rational addiction model. Gruber and Köszegi (2001) also find evidence supporting forward-looking behavior but argue that individual preferences may be time-inconsistent. Arcidiacono, Sieg, and Sloan (2007) develop a different test of the rational addiction model. Rather than seeing if substance use responds to prices, they explore whether substance use of the near-elderly responds to negative health and income shocks. They similarly find that forward-looking models fit the data better than myopic models. This literature motivates my decision to model individuals as rational, forward-looking decision makers.

Relationships Between Substances: There are several papers that study whether cigarettes, alcohol, and marijuana are substitutes or complements among youth. Dee (1999) studies the relationship between alcohol and cigarette use of youth using the Monitoring the Future surveys of high school seniors. He utilizes changes in cigarette taxes and the state minimum legal drinking ages and finds that increasing the minimum legal drinking age decreases teen smoking and that higher cigarette taxes are associated with lower alcohol use. Therefore, he concludes that alcohol and cigarettes are economic complements. Krauss et al. (2014) also find that cigarette and alcohol are complements and suggest increasing the tax on cigarettes in order to lower alcohol consumption. Farrelly et al. (2001) and Chaloupka et al. (1999) study the relationship between cigarettes and marijuana. Farrelly et al. (2001) use the National Household Survey on Drug Abuse and find that higher cigarette taxes decrease the intensity of marijuana use, suggesting that cigarettes and marijuana are complements. Chaloupka et al. (1999) use data from the Monitoring the Future Project to similarly show that higher cigarette prices reduce the level of marijuana used by current users.

There is no clear consensus in the literature about whether or not alcohol and marijuana are complements or substitutes. DiNardo and Lemieux (2001) use state aggregated data from the Monitoring the Future Survey to study the relationship between alcohol and marijuana consumption of youth. They find that increases in the minimum legal drinking age reduce the prevalence of alcohol use, but increase the prevalence of marijuana use. Thus, they conclude that alcohol and marijuana are substitutes. Similarly, Chaloupka and Laixuthai (1997) use micro-level data from Monitoring the Future Survey to estimate the demand of alcohol, where alcohol use is a function of alcohol prices, legal drinking ages, and the price of marijuana. They find that marijuana decriminalization laws decrease alcohol prevalence, suggesting that marijuana and alcohol are substitutes. Thies and Register (1993) and Pacula (1998) use individual level data from the National Longitudinal Survey of Youth to jointly estimate demand equations for both marijuana and alcohol. Thies and Register (1993) find that marijuana decriminalization and the minimum legal drinking age have no statistically significant effect on alcohol or marijuana use. Pacula (1998) finds that increases in the beer tax decrease marijuana use, suggesting that marijuana and alcohol are complements. Williams et al. (2004) study the relationship between alcohol and marijuana use among college students using the Harvard School of Public Health's College Alcohol Study. They also find that alcohol and marijuana are complements. By jointly modeling the choices of cigarettes, alcohol, and marijuana, I am able to estimate whether or not these substances are substitutes or complements among youth.

Substance Use Effects on Educational Outcomes: Several authors have studied the effects of substance use on educational outcomes such as high school graduation and college enrollment. There is no clear consensus on how substance use affects educational attainment. Cook and Moore (1993) find that students who attend high school in states with high beer taxes or a high minimum drinking age complete more years of schooling. Yamada, Kendix, and Yamada (1996) find significant adverse effects of current alcohol and marijuana use on high school graduation rates. Alternatively, Dee and Evans (2003) conclude that alcohol consumption of teenagers does not reduce educational attainment. Bray et al. (2000) and Chatteriji (2006) find evidence that marijuana use decreases the probability of high school graduation.

In my model, I am not relying solely on variation in state laws or prices to determine an individual's substance use. Individuals are also influenced by shocks that affect their educational attainment, wages, and job opportunities. I also model the usage of several substances, allowing me to control for the interdependencies of cigarette, alcohol, and marijuana use. By explicitly modeling substance use decisions within a human capital accumulation model, I can determine the channel through which substance use affects educational attainment. That is, I can determine whether lower achievement among users is due to decreased enrollment, increased probability of failing, or because individuals who use substances simply do not enjoy school. Lastly, the estimates of my model allow me to conduct counterfactual simulations, for example, of the effect of changes in substance legality on educational attainment and employment decisions.

Substance Use Effects on Labor Market Outcomes: Several studies have examined the effects of substance use on labor market outcomes, such as wages, employment, and hours of work. Often, authors use past substance use, rather than current substance use, in their regressions, which they argue is independent of the error term (Buchmueller and Zuvekas 1998, Zarkin et al. 1998). DeSimone (2002), on the other hand, instruments for current substance use and finds that the use of marijuana and cocaine each substantially reduces the likelihood of employment. Kaestner (1994) also treats substance use as endogenous and, after controlling for unobserved heterogeneity, he finds no effect of illicit drug use on labor supply. All of the papers listed above treat education as an exogenous determinant of labor market outcomes.

Bray (2005) improves on this earlier work by treating alcohol use, educational attainment, and work experience as endogenous. In his model, he allows educational attainment, work experience, and alcohol use to affect wages. He models the enrollment, work, and alcohol use decisions as logit models that are designed to capture the correlation of the decisions with the error term in the wage equation. The reduced-form equations are auxiliary estimating equations and, therefore, cannot be directly interpreted. Estimation results suggest that moderate alcohol use while in school or working has a positive effect on human capital accumulation, and heavier drinking reduces this gain. My model differs from

Bray's in two ways. First, I develop and estimate a structural model of individuals' schooling, work, and substance use decisions, which allows me to see how current substance use affects not only wages but also schooling decisions and future substance use. Second, individuals in my model make decisions about alcohol, cigarette, and marijuana use, which allows me to account for the interdependencies of substance use.

2.2 Human Capital Investment Literature

Most papers on human capital investment are based on Roy (1951) and Heckman and Honore (1990), which investigate how selection into occupations affects the distribution of earnings and productivity in those occupations. Eckstein and Wolpin (1999) use this framework to look at youths' high school dropout decisions. Their findings indicate that dropping out of high school is confined to youths with lower ability, a lower expected value of a high school diploma, a comparative advantage in skills suited for jobs that do not require a high school diploma, and a lower consumption value of attending school. Whereas Eckstein and Wolpin (1999) mainly focuses on educational outcomes, another set of papers focuses more on labor market outcomes. Keane and Wolpin (1997) and Sullivan (2010) examine dynamic educational and occupational choices and allow work experience and education to be accumulated endogenously. In Keane and Wolpin (1997), individuals choose whether to attend school, to work in one of three occupations, or to engage in home production. Sullivan (2010) expands this choice set by allowing individuals to select into more than three occupations, to select into firms in addition to occupations, and to participate in dual activities such as employment while attending school. Following Sullivan (2010), I allow individuals to select into firms, and I allow for dual activities. Unlike Sullivan (2010) and

Keane and Wolpin (1997), I allow individuals to pick part-time or full-time hours, which is particularly important for individuals while they are enrolled in school.

Mezza (2011) is most closely related to my paper. He estimates a dynamic structural model where individuals jointly make decisions about whether to consume drugs, attend school, and participate in the labor force. He finds that non-drug users have higher wages than marijuana and/or hard drug users. My model differs from the model presented in Mezza (2011) in three main ways. Most importantly, I allow individuals to choose different levels of substance use. Secondly, I allow substance use to affect the probability that an individual gets arrested and allow arrests to affect human capital accumulation choices. This allows me to determine whether it is the use itself that is affecting outcomes or if the effect is coming through arrests. Lastly, my focus is on softer drugs, that is cigarette, alcohol, and marijuana use, whereas his focus is on harder substance use. Cigarette and alcohol use may have strong gateway effects among youth that Mezza (2011) is not able to capture. These gateway effects could have strong policy implications for combating youth substance abuse.

3 Model

In this section, I show how young men make decisions about schooling, employment, and substance use. Some modeling choices, especially about substance use, are heavily influenced by data availability, which is discussed in more detail in Section 4. I use a finite horizon, discrete time, dynamic model. In each year, from age 14 ($t = 1$) to a known terminal age ($t = T$), an individual i makes decisions in order to maximize the discounted sum of his lifetime utility subject to annual budget constraints. An individual makes choices about his substance use as well as about his employment and education. He chooses whether to consume no cigarettes, a moderate amount of cigarettes, or a heavy amount of cigarettes, $cigs_{it} \in [1, 2, 3]$; how many days to consume alcohol (none, moderate, or heavy), $alc_{it} \in [1, 2, 3]$; and how many days to consume marijuana (none, moderate, or heavy), $mj_{it} \in [1, 2, 3]$. Let $sub_{it} = (cigs_{it},\ alc_{it},\ mj_{it})$ denote the substance use choices of individual i at age t. The human capital accumulation choices individuals make are: whether to enroll in school, $enroll_{it} = 1$; whether to be unemployed $h_{it} = 0$, work part-time $h_{it} = 1$, or work full-time $h_{it} = 2$; and, if working, whether to work for a new employer, $ne_{it} = 1$. I denote the vector of human capital accumulation choices as $hc_{it} = (enroll_{it}, h_{it}, ne_{it})$. An individual receives one new part-time offer and one new full-time offer each period. If he worked in the previous period, then he also has the choice to continue working at the same job. An individual can always choose not to work or to enroll in school. Let $K_{it}\left(s_{it}\right)$ denote the set of human capital accumulation choices an individual faces, where s_{it} denotes the state that individual i is in at time t. The choice set that individual i faces, $K_{it}\left(s_{it}\right)$, varies with s_{it} because individuals who worked last period have an additional discrete choice, whether to continue working at the same job .

Individuals are observed choosing different paths due to observed and unobserved heterogeneity that affect the utility associated with each choice. For example, someone may choose to work at a very low wage while enrolled in school because the experience and education will raise their wages in the future, while someone who has a lower preference for work might choose to not work. My model allows substance use to affect career paths through its effect on educational attainment, wage offers, and arrests.

The effect of use on educational attainment is ambiguous, operating through enrollment and grade completion. School may make obtaining drugs easier, especially if other youths in the school are using drugs, which would encourage users to enroll in school. However, substance use could make it harder for a youth to succeed in school, as using substances may crowd out time spent on schoolwork or decrease cognitive ability. These effects raise the cost for a person using substances to put forth the same amount of effort towards school as a person not using and may discourage enrollment. Individuals in my model can only choose whether or not to enroll, not their educational attainment. Individuals who choose to enroll have a probability of passing that is a function of their substance use.

I also allow substance use to affect work choices through two channels. First, substance use may affect whether or not an individual chooses to work. For example, an individual using a large amount of alcohol may find it costly to work everyday in comparison to someone not using alcohol. Current substance use can also affect job opportunities by altering productivity and hence the wage offers the individual receives. Lastly, I allow substance use to affect whether or not an individual gets arrested. I allow arrests to affect the enrollment and work decisions as well as the wage offers.

3.1 Choice-Specific Utility Flows

In each period t, utility is comprised of pecuniary utility from the consumption of goods $CONS_{it}$ and nonpecuniary utility N_{it}. The utility an individual gets from making choice (hc_{it}, sub_{it}) is a function of endogenous state variables s_{it}, skill and preference endowments, and random shocks that vary over time, people, and employers. The endogenous state variables measure educational attainment, work experience, firm-specific human capital, arrest histories, and the addictive stocks of the three substances. The utility flow that individual i receives at time t from choices hc_{it} and sub_{it} is

$$u_{it}\left(hc_{it}, sub_{it}|s_{it}\right) = \ln\left(CONS_{it}\right) + \ln\left(N_{it}\right) \tag{1}$$

The remainder of this subsection describes the structure of the pecuniary and nonpecuniary utility flows.

3.1.1 Pecuniary Utility Flows

Assuming no savings, an individual's budget constraint equates his expenditures $CONS_{it}$ to his income.[2] Potential earnings from working h_{it} hours at firm k is denoted w^h_{itk}. It is not straightforward to measure and model income, consumption, and wealth of youths receiving support from their parents. Yet, youths make substance use and work decisions that take income into account. Therefore, I allow individuals to receive transfers $\widetilde{w}_{it}\left(hc_{it}\right)$ from an outside source, for example, family members or the government.[3] Expenditures include purchases of a composite commodity with price equal to one; purchases of cigarettes at price p^c_{it}, alcohol at price p^a_{it}, and marijuana at price p^m_{it}; and expenses p^s_{it} associated with attending college. Thus, in each period, an individual's budget constraint is

$$CONS_{it} + p^{sub}_{it} sub_{it} + p^s_{it} \mathbb{1}\left(G_{it} > 12\right) \mathbb{1}\left(enroll_{it} = 1\right) = w^h_{it} h_{it} + \widetilde{w}_{it} \tag{2}$$

where $p^{sub}_{it} = \left(p^c_{it,} \, p^a_{it}, \, p^m_{it}\right)$, $\mathbb{1}\left(\cdot\right)$ is an indicator that equals one if the argument is true, and G_{it} is the educational attainment at time period t. I assume transfers are large enough so that consumption is above a minimum level, C_{min}.

[2] I rule out saving for college or to purchase substances as a motivation for working while in high school. Modeling savings would require using accumulated earnings as a state variable that affects school and substance use decisions.

[3] Similarly to Eckstein and Wolpin (1999), I do not model the parents' decision about how much to transfer to their children, so $\widetilde{w}_{it}$ is treated as exogenous. I allow the distribution of transfers to vary by discrete choice to capture the fact that an individual not working and enrolled in college needs more (and may be rewarded with more) transfers than an individual not going to school and working full-time. This is to avoid having to model game-theoretic interactions between youths and their parents, as in Rosenzweig and Wolpin (1994); Hao, Hotz, and Jin (2008); and Martinez (2014)

Individuals who work at the same firm for several years are able to build firm-specific human capital that raises their productivity. Individuals in my model are allowed to choose to work at the firm they worked at in the previous period or at a new firm. An individual's log-wage for hours $h_{it} = h$ at firm k in period t is

$$\ln\left(w_{itk}^{h}\right) = \theta_{w}^{h}\left(sub_{it}, s_{it}\right) + \psi_{ik} + \mu_{i}^{w} + \nu_{itk}^{h}. \tag{3}$$

The term $\theta_{w}^{h}\left(sub_{it}, s_{it}\right)$ represents the deterministic part of the log-wage, which is a function of human capital, arrest history, and today's substance use. This specification allows the effects of state variables and substance use on the wage to differ by hours choice. ψ_{ik} is a permanent worker-firm productivity match value. This reflects factors, such as an individual's rapport with his boss, which affect the wage of worker i at firm k. Time-persistent individual heterogeneity in productivity that does not vary over time nor jobs is denoted μ_{i}^{w}. The last term ν_{itk}^{h} is an idiosyncratic shock that captures true randomness in wages. Individuals observe all of the components of the wage when a job offer is received. Future realizations of firm-specific match values and wage shocks are unknown to the individual until they occur.

3.1.2 Nonpecuniary Utility Flows

Nonpecuniary utility is a function of individual characteristics and state variables. Numerous forms of state dependence may affect the ease or difficulty with which one can change states; therefore, nonpecuniary utility may differ according to one's choices in the previous period. For example, it is more difficult for a person to go to school after being unenrolled. Also, tenure in a state may affect nonpecuniary utility. For example, it may be costly for an individual who has worked at the same firm for several periods to move to a new firm. Allowing tenure in substance use states to enter into the utility function also allows me to measure the addictiveness of substances. The log-nonpecuniary utility flow in period t is

$$\begin{aligned} \ln(N_{it}) &= n(hc_{it}, sub_{it}, s_{it}) \\ &+ \mu_i^{sub} sub_{it} + \mu_i^h \mathbb{1}(h_{it} = h) + \mu_i^s \mathbb{1}(enroll_{it} = 1) \\ &+ \epsilon_{it}^{hc} + \epsilon_{it}^{sub} \end{aligned} \tag{4}$$

where $\mu_i^{sub} = \left(\mu_i^{cigs=2}, \mu_i^{cigs=3}, \mu_i^{alc=2}, \mu_i^{alc=3}, \mu_i^{mj=2}, \mu_i^{mj=3}\right)$. The first line in Equation 4 represents the deterministic part of the nonpecuniary utility that is a function of the choices and the state vector. This term includes the nonpecuniary utility flows an individual receives from employment, unemployment, and substance use as well as the cost function of attending school. The deterministic part of nonpecuniary utility is

$$\begin{aligned} n(hc_{it}, sub_{it}, s_{it}) &= \sum_{h=0}^{2} \theta_N^h(sub_{it}, s_{it}) \mathbb{1}(h_{it} = h) \\ &+ \kappa_S(h_{it}, sub_{it}, s_{it}) \mathbb{1}(enroll_{it} = 1) \\ &+ \alpha(sub_{it}, s_{it}). \end{aligned} \tag{5}$$

The first line of Equation 5 captures the nonpecuniary utility of employment. This specification allows the effect of state variables and substance use on employment utility to vary by hours choice h. For example, an individual may get more enjoyment out of working full-time relative to not working if he is not consuming marijuana. This term also captures the utility gained from the leisure time consumed when unemployed. The second line is the nonpecuniary cost of enrolling in school. The cost function is allowed to differ depending on employment and substance use because it may be more difficult to attend school while employed full-time rather than part-time or when drinking heavily. The third line captures the utility gained from using substances.

The second line in the non-pecuniary utility in Equation 4 captures time-persistent unobserved heterogeneity in preferences for substance use, hours worked, and school enrollment. Including this source of heterogeneity may help explain why some individuals always drink alcohol but never smoke cigarettes. It also allows for heterogeneity in the cost of working and schooling caused by unobserved characteristics such as ability. The final line in Equation 4 is an idiosyncratic shock to the nonpecuniary utility flow that person i receives at time t from making the choice (hc_{it}, sub_{it}).

3.2 State Variables

The state variables in the vector s_{it} can be divided into continuous state variables s_{it}^c and discrete state variables s_{it}^d. The continuous state variables are the addictive stocks of the substances. The discrete state variables include measures of human capital as well as the number of arrests.

I allow for the accumulation of addictive stocks in each substance, which captures the potential costs of changing substance use from one period to the next. These are the addictive stock associated with cigarette use ($\mathscr{C}_{it}$); with alcohol use ($\mathscr{A}_{it}$); and with marijuana use ($\mathscr{M}_{it}$). The addictive stock associated with cigarettes evolves as follows:

$$\mathscr{C}_{it} = \begin{cases} \beta_1^c \mathscr{C}_{i,t-1} & if\ cigs_{i,t-1} = 1 \\ \mathscr{C}_{i,t-1} + \beta_2^c & if\ cigs_{i,t-1} = 2 \\ \mathscr{C}_{i,t-1} + 1 & if\ cigs_{i,t-1} = 3 \end{cases}. \tag{6}$$

If an individual chose not to consume cigarettes in period $t-1$, $cigs_{i,t-1} = 1$, then his addictive stock in period t will depreciate by $\beta_1^c \in (0,1)$. If instead an individual chose to consume a moderate amount, $cigs_{i,t-1} = 2$, then his addictive stock will increase by $\beta_2^c \in (0,1)$. $\mathscr{A}_{it}$ and $\mathscr{M}_{it}$ evolve in the same way. β_1^c and β_2^c are important, because they allow me to capture certain patterns of use, such as it being easier for an individual to not use a substance in the current period if he did not use the substance in the previous period. Given this specification, the continuous state variables are $s_{it}^c = (\mathscr{C}_{it}, \mathscr{A}_{it}, \mathscr{M}_{it})$.

Human capital is measured by educational attainment and work experience. Educational attainment is summarized by the years of schooling G_{it} an individual has completed, where

$$G_{it} = G_{i,t-1} + g_{it}. \tag{7}$$

$g_{it} = 1$ if an individual completes a year of schooling and zero otherwise. A student who enrolls in school will not necessarily pass that grade. If an individual enrolls in school, the probability that he completes the grade, π_{it}^{school} , is exogenously determined by the individual's states and choices. That is

$$\pi_{it}^{school} = \Phi\left(\eta_{school}\left(sub_{it}, s_{it}\right) + \mu_i^{pass}\right), \tag{8}$$

where, $\Phi(\cdot)$ is a standard normal cdf, η_{school} is the deterministic part of the probability, and μ_i^{pass} represents the individual's unobserved ability in school. η_{school} is a function of substance use, educational attainment, whether the individual enrolled last period, and whether he passed last period.

Work experience is measured by total hours H_{it} worked at any firm and by firm-specific human capital τ_{it}. I only keep track of the firm-specific human capital at the most recent in order to simplify the state space. That is,

$$H_{it} = \begin{cases} H_{i,t-1} & if\ h_{it} = 0 \\ H_{i,t-1} & with\ prob\ 1 - \pi^H\ if\ h_{it} = 1 \\ H_{i,t-1} + 1 & with\ prob\ \pi^H\ if\ h_{it} = 1 \\ H_{i,t-1} + 1 & if\ h_{it} = 2 \end{cases}, \tag{9}$$

where π^H is the probability that a part-time job increases human capital by one unit. This probability is estimated along with the other parameters of the model.[4] τ_{it} evolve as

$$\tau_{it} = \begin{cases} 0 & if\ h_{it} = 0\ or\ ne_{it} = 1 \\ \tau_{i,t-1} + 1 & if\ h_{it} > 0\ and\ ne_{it} = 0 \end{cases}. \tag{10}$$

So, tenure goes up by one if an individual decides to work at the same firm he worked at last year; that is, $h_{it} > 0$ and $ne_{it} = 0$. Tenure goes to zero if he chooses not to work or to work at a new firm.

[4]This is a simplification used to decrease the size of my state space.

Lastly, if an individual gets arrested in period t, then $r_{it} = 1$. Therefore, the number of arrests evolves as

$$R_{it} = R_{i,t-1} + r_{it}. \tag{11}$$

The deterministic part of the probability of getting arrested, η_{arrest}, is a function of substance use, previous arrest history, age, and local law enforcement characteristics. The probability is also a function of an individual's unobserved propensity to get arrested, μ_i^{arrest}. The probability of arrest is written as

$$\pi_{it}^{arrest} = \Phi\left(\eta_{arrest}\left(sub_{it}, s_{it}\right) + \mu_i^{arrest}\right) \tag{12}$$

The discrete state variables are $s_{it}^d = (G_{it}, g_{it-1}, enroll_{i,t-1}, R_{it}, H_{it}, \tau_{it}, h_{i,t-1})$.

3.3 The Individual Optimization Problem

I assume individuals are forward-looking. Each period an individual makes choices in order to maximize his present discounted value of expected lifetime utility subject to the budget constraint in Equation 2. At the beginning of the first period, the individual knows the wage function of each hours choice and the deterministic components of the utility function. He also knows his skill endowment μ_i^w and his choice-specific preference values $\mu_i^N = \left(\mu_i^{sub}, \mu_i^h, \mu_i^s, \mu_i^{arrest}, \mu_i^{pass}\right)$. Let $\mu_i = \left(\mu_i^w, \mu_i^N\right)$. Lastly, the individual knows his current levels of the addictive stocks $\mathscr{C}_{i0}$, $\mathscr{A}_{i0}$, and $\mathscr{M}_{i0}$. Future realizations of firm-specific match values ψ_{ik}, wage shocks ν_{ikt}^1 and ν_{ikt}^2, and the choice-specific utility shocks corresponding to the human capital accumulation choices ϵ_{it}^{hc}, and to the substance use choices ϵ_{it}^{sub}, are unknown to the individual until they occur; however, he knows the distributions of these variables. He also knows the probabilities that he will get arrested and will complete a year of schooling conditional on enrollment.

The optimization problem can be represented in terms of choice-specific value functions which give the lifetime discounted value of each choice for a given set of state variables, s_{it}.

The value function for individual i in period t is

$$\begin{aligned} V\left(s_{it}\right) &= \max_{hc_{it}\in K_{it}(s_{it}),sub_{it}} v_{it}\left(hc_{it}, sub_{it}|s_{it}\right) \\ s.t. \quad & CONS_{it} + p^{sub}_{it} sub_{it} + p^{s}_{it} \mathbb{1}\left(G_{it} > 12\right) \mathbb{1}\left(enroll_{it} = 1\right) = w^{h}_{it} h_{it} + \widetilde{w}_{it} \\ & CONS_{it} \geq C_{min} \end{aligned} \tag{13}$$

where the choice set that individual i faces, $K_{it}\left(s_{it}\right)$, varies with s_{it} because individuals who worked last period have an additional discrete choice, whether to continue working at the same job. The term $v_{it}\left(hc_{it}, sub_{it}|s_{it}\right)$ in Equation 13 is the choice-specific value function

$$v_{it}\left(hc_{it}, sub_{it}|s_{it}\right) = u_{it}\left(hc_{it}, sub_{it}|s_{it}\right) + \delta EV\left(s_{i,t+1}|s_{it}, hc_{it}, sub_{it}\right), \tag{14}$$

where δ is the discount factor. $EV\left(s_{i,t+1}|s_{it}, hc_{it}, sub_{it}\right)$ is the expected value function in period $t+1$ conditional on states and choices made in period t. Expectations are taken over the random shocks to utility and wages, future match values at new jobs, future arrests, and future educational attainment. For an individual who is not enrolled in school,

$$EV\left(s_{i,t+1}|s_{it},hc_{it},sub_{it}\right) = \pi_{it}^{arrest}EV\left(s_{i,t+1}|s_{it},hc_{it},sub_{it},r_{it}=1\right) + \left(1-\pi_{it}^{arrest}\right)EV\left(s_{i,t+1}|s_{it},hc_{it},sub_{it},r_{it}=0\right) \quad (15)$$

where π_{it}^{arrest} is the probability that the individual gets arrested. At time t, the individual does not know if he will get arrested, $r_{it} = 1$, but he knows π_{it}^{arrest}. Therefore, his expectations are taken over the probability of arrest. If someone is enrolled in school, then the expected value function includes the probability that the individual passes that grade, $g_{it} = 1$. That is,

$$\begin{aligned} EV\left(s_{i,t+1}|s_{it},hc_{it},sub_{it}\right) &= \pi_{it}^{school}\pi_{it}^{arrest}EV\left(s_{i,t+1}|s_{it},hc_{it},sub_{it},g_{it}=1,r_{it}=1\right) \\ &+\pi_{it}^{school}\left(1-\pi_{it}^{arrest}\right)EV\left(s_{i,t+1}|s_{it},hc_{it},sub_{it},g_{it}=1,r_{it}=0\right) \\ &+\left(1-\pi_{it}^{school}\right)\pi_{it}^{arrest}EV\left(s_{i,t+1}|s_{it},hc_{it},sub_{it},g_{it}=0,r_{it}=1\right) \\ &+\left(1-\pi_{it}^{school}\right)\left(1-\pi_{it}^{arrest}\right)EV\left(s_{i,t+1}|s_{it},hc_{it},sub_{it},g_{it}=0,r_{it}=0\right). \end{aligned}$$

If someone works part-time, then the expected value will similarly incorporate the probability that human capital, $H_{i,t+1}$, increases by one. Individuals use the value functions to determine the optimal educational, employment, and substance use choices each period.

3.4 Identification

In this subsection, I will discuss the identification of the causal effects of substance use on outcomes as well as the identification of the parameters in the model. The causal effects of substance use on education and career outcomes are difficult to identify since individuals endogenously select into substance use, enrollment, and employment states. One concern is that, for example, those who are more likely to use large amounts of marijuana are those for whom the labor market would offer lower wages regardless of marijuana use, other things held constant. Alternatively, it may be the case that the only individuals using marijuana are those who will not be negatively affected by it. In this case, ignoring unobservable characteristics that are important in both selection into marijuana use and wages will overstate the negative effects of marijuana use on wages. In order to identify the effects of substance use on outcomes, I need selection into enrollment, substance use, and employment states to each be partly explained by exogenous variation that is not correlated with the outcome of interest. I use the cost of higher education in the respondent's state of residence as an exclusion restriction to help explain educational choices; the prices of substances to help explain substance use; local labor market conditions to help explain employment outcomes; and local law enforcement characteristics to help explain arrests.[5]

Next, I will discuss the identification of the parameters in the model. Individuals are observed making different choices because they differ in characteristics that affect their opportunities, for example through wages. The parameters of the wage equations are identified by the co-variation in wages and observable individual characteristics. I also assume that individuals have unobserved characteristics like ability that affect their wages. Based on observables, I can predict a person's wage. An individual with high unobserved ability will have persistently higher than predicted wages. This identifies the distributional parameters of individual heterogeneity in productivity. Some individuals may have higher-than-predicted wages because they have a particularly high worker-firm match value, perhaps because they get along with their boss; if such people change jobs, they may no longer have a higher than predicted wage. The parameters of the distribution of the worker-firm

[5] I show in the appendix that substance prices affect whether or not 14 year olds start using substances and local unemployment rates affect whether or not an individual is working.

match value are identified by the correlation in the residuals of wages within a job after controlling for persistent personal effects that arise across jobs. The distributional parameters of the wage shock are identified by the variation in wages across individuals conditional on individual characteristics and worker-firm match values, arising because wages suddenly rise, for example, for a worker who appears to have low ability and low match value.

Individuals do not necessarily choose the option that maximizes pecuniary utility. The covariation in observed individual characteristics and choices that are not explained by maximizing pecuniary utility identify the parameters of the nonpecuniary utility. Individuals who are observationally equivalent make different work, schooling, and substance use decisions due to heterogeneity in time-persistent preferences. Based on observables, I can predict what the optimal choice is for an individual. The correlation in residuals between the predicted and observed choices within an individual identifies the distributional parameters of the unobserved tastes that affect nonpecuniary utility.

4 Data

4.1 NLSY97

I use the 1997 National Longitudinal Survey of Youth (NLSY97), which consists of 4,599 male youths who were 12 to 16 years old as of December 31, 1996. Interviews have been conducted annually since 1997; I use the first 13 waves of data, until they were aged 24-28 in 2009. The NLSY97 collects information about labor market behavior, educational experiences, and drug use. The NLSY97 consists of a nationally representative core sample and a supplement that over-samples blacks and Hispanics. I use the entire sample in my analysis. I limit my sample to individuals whom I observe at age 14, which is the first age at which individuals are asked about their substance use. An individual remains in the data set until the observation is truncated at the first instance of missing information for any variable that I use. My sample consists of 1,170 individuals whom I observe on average for 7.65 years, providing 8,946 person-year observations. Table 1 shows how I obtain my sample. The biggest decrease in my sample size comes from dropping individuals who are not interviewed at age 14. I assume that all individuals have not used substances prior to age 14. Therefore, their addictive stock is zero at age 14. If I do not observe a person starting at age 14, then I do not know what their substance use stock is.[6]

Table 1: Sample Selection

Rule	Individuals Lost
Missing Substance Use or Arrest Information at Age 14	185
Missing Educational Attainment at Age 14	3
Missing Location at Age 14	6
Race Is Mixed	16
Not Interviewed at Age 14	2,270
Only Observed at Age 14	949
Final Sample Size	1,170

Data come from the NLSY97.

The original sample consists of 4,599 male youths who were 12 to 16 years old as of December 31, 1996.

The decision period of my model corresponds to a school year, which runs from August to July. I use monthly school enrollment arrays to construct enrollment status. My model allows individuals to fail a grade; therefore, I consider an individual as enrolled in a particular school year if he reports being enrolled for at least 1 month of that school year. Among those enrolling in grades K-12, 10.96 percent of my sample fails to advance to the

[6]At a future date, I plan to adapt my estimation strategy to simulate this missing information, which will allow me to use more individuals. Data augmentation methods provide a simple way to do this within my Bayesian MCMC estimation strategy.

next grade, whereas 29.87 percent of those enrolled in college fail.[7] The amount of education accumulated is determined using a variable that indicates the highest grade completed as of the interview date.

The NLSY97 has week-by-week reports of an individual's working status. I use weekly hours of work in a school year to determine annual hours. I classify an individual as working full-time if he works at least 30 hours per week and 45 weeks per year, or at least 1,350 hours per year. To determine the relevant wage for each working individual, I sum up the hours worked each week during the school year for each job. The wage rate for that school year is the wage at the job where the individual worked the most hours. Annual earnings is defined as that wage multiplied by 2,000 hours for full-time workers and 1,000 for part-time workers.[8] The NLSY97 offers two measures of hourly wages. The first is the hourly wage where overtime and performance pay is excluded. The second includes all extra compensation such as over-time, tips, and bonuses. I choose the second measure because it gives me a better measure of the income individuals have to spend on consumption. Additionally, in my model, I assume that the wage reflects the marginal productivity of the individual, and it is reasonable to assume that increases in performance pay indicate higher productivity, making the measure that includes tips and bonuses more appropriate.

In each wave, individuals are asked questions about their use of cigarettes, alcohol, and marijuana in the past 30 days. I assume that this information is representative of their use over the entire school year.[9] Since substance use is a sensitive topic, these questions are administered through the use of audio computer-assisted self-interview technology rather

[7] According to Acaldi et al. (2011), the median time to graduation for 2008 bachelor's degree recipients to earn their degree was between 52 and 80 months, depending on their path to graduation.

[8] Restricting the choice of hours worked to either 1,000 or 2,000 hours is done for tractability.

[9] I compare the self-reported substance use rates in the NLSY97 with those in the 1997 National Household Survey on Drug Abuse (NHSDA) in Table 16 in the appendix. Rates are slightly higher in the NLSY97, but overall are comparable to those in the NHSDA.

than an interviewer, so I treat the answers as truthful. Harrison at al. (2007) report on a validity study of self-reported substance use. The study was conducted in conjunction with the National Household Survey on Drug Abuse, which uses the same technology as the NLSY97. Urine samples were collected for a subset of respondents in order to compare self-reported substance use with actual substance use. For tobacco, 88.7 percent of individual's self-reported use in the past 3 days agreed with their urine sample; 7.7 percent reported no use and tested positive for use while 3.6 reported using and tested negative. For marijuana, 93 percent of individual's self-reported use in the past 3 days agreed with their urine sample; 5.2 percent reported no use and tested positive and 1.8 percent reported using and tested negative.

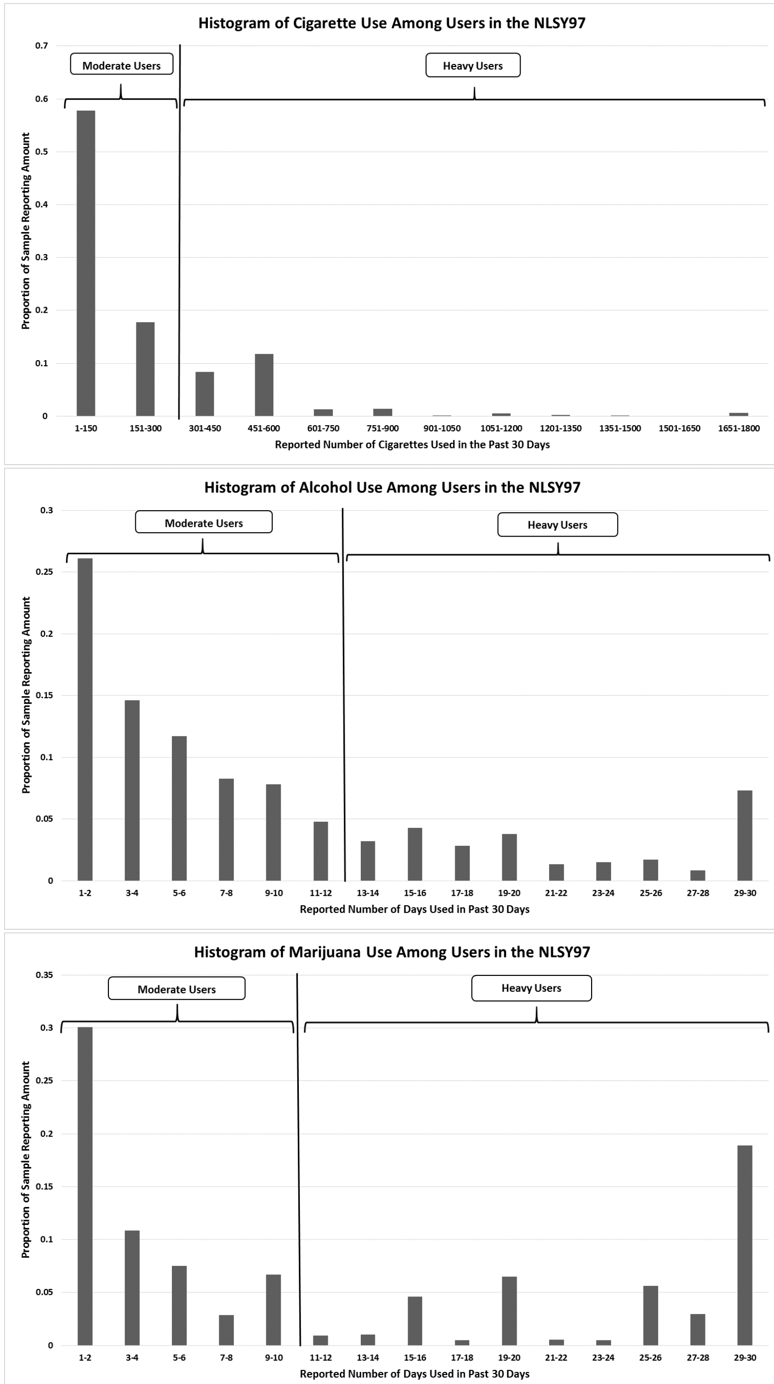
Individuals in my model choose how many cigarettes to consume, how many days to consume alcohol, and how many days to consume marijuana. In particular, I allow individuals to choose whether to consume none of a substance, a moderate amount, or a heavy amount. I classify moderate cigarette use as consuming more than zero but less than 300 cigarettes per 30 days; moderate alcohol use as more than zero but less than 13 days; and moderate marijuana use as more than zero but less than 11 days. The amount consumed in each category is determined by the mean consumption of that substance observed in the data. Figure 1 shows the distribution of substance use among those using a positive amount as well as how I classify moderate and heavy use.

4.2 Price Data

I use the cigarette prices reported in Volume 46 of The Tax Burden on Tobacco (Orzechowski and Walker 2011).[10] It includes federal taxes, state taxes, and the average retail

[10]This publication can be found at http://www.taxadmin.org/fta/tobacco/papers/Tax_Burden_2011.pdf.

Figure 1: Histograms of Substance Use Among Users in the NLSY97



Data are from the NLSY97. The sample used is described in Table 1.

price for a pack of cigarettes from 1955-2010. Table 2 presents summary statistics of the total price of a pack of cigarettes by year, which I define as the average retail price plus taxes. I show that the average price of cigarettes increases significantly over my time span. It rises from $2.64 per pack in 1997 to $4.52 per pack in 2008. In addition, there is significant variation across states.

As is common in the literature, I use the Cost of Living Index collected by the American Chamber of Commerce Research Association (ACCRA) for alcohol prices.[11] The index is reported quarterly and collects prices from around 300 cities for a 6-pack of beer. I assume that individuals living in the same state face the same prices, so I aggregate prices to state-year observations. I first average each city's price across quarters to get a city-year price. Then, to get a state price, I calculate the weighted average where weights are proportional to an included city's population.[12] The Cost of Living Index does not collect data from every city each quarter, so I am still left with some states that do not have a price during certain years. I impute these missing prices using the regression

$$P_{sy} = \beta_1 + \beta_2 statetax_{sy} + \beta_3 \mathbb{1}\left(year_y > 2000\right) + \alpha_s + \eta_y + \epsilon_{sy},$$

where P_{sy} is the beer price in state s and year y, $statetax_{sy}$ is the state beer tax, α_s is the state fixed effect, and η_y is the year fixed effect. Prior to 2000, ACCRA collected the price of a 6-pack of Budweiser or Miller, and then switched to Heineken in 2000. Therefore, I include $\mathbb{1}\left(year_y > 2000\right)$ as a regressor to adjust for the change in the type of beer sampled after 2000. The regression produces an R-squared of .97. I then impute

[11] Ruhm et. al. (2012) find that "barcode" scanner data, collected by AC Nielsen provide a better price elasticity of demand for beer than the ACCRA data. While I would prefer to use scanner data, the data only started being collected in 2006 and only recently became accessible to researchers.

[12] The populations were collected from the U.S. Census Bureau.

the missing values using the estimated coefficients and subtract β_3 from observations after the year 2000 to account for the sampling change. Regression results are presented in the appendix. I present summary statistics of the prices I use in Table 2. The average price of beer decreases over time from $4.65 for a 6-pack in 1997 to $4.15 in 2008.

I use marijuana price data assembled from High Times magazine.[13] Most papers in the literature use marijuana prices from the Drug Enforcement Administration's (DEA) System to Retrieve Information from Drug Evidence (STRIDE). The DEA's focus is on harder drugs like heroin, so the number of marijuana observations are small, and a large amount of those observations are concentrated in Washington D.C. The data collected from High Times has better geographic coverage.

In each monthly issue, contributors to the "Trans High Market Quotations" (THMQ) section write in with a description of the marijuana, the price per ounce, and the city and state they live in. Using the description of the marijuana, I divide observations into what I am calling low-grade marijuana and high-grade marijuana, where high-grade marijuana has a higher potency of THC, the active drug in marijuana. To divide the observations into grades, I found nicknames for low-grade marijuana on marijuana forums and classified an observation as low-grade if it includes keywords such as: Schwag, Brick, Dirt, Mids, Commercial. More details can be found in Alford (2013). I aggregate both the low-grade and high-grade prices to the state-year level in the same way as I do with the alcohol prices. I then estimate the following regression to impute missing state-year prices,

[13] Mireille Jacobson provided me data from 1996-2005, and Mark Anderson provided me with data from 2008-2010. I collected data from 2006-2007 online at www.hightimes.com. See Jacobson (2004) and Anderson, Hansen, and Rees (2013) for more information about the price data.

$$\begin{aligned} P_{sy} &= \beta_1 + \beta_2 medicalmjlaw_{sy} + \beta_3 decriminalized_{sy} + \beta_4 violentcrimerate_{sy} \\ &+ \beta_5 murderrate_{sy} + \beta_6 propertycrimerate_{sy} + \beta_7 lowgrade \\ &+ \alpha_s + \eta_y + \epsilon_{sy}, \end{aligned}$$

where $medicalmjlaw_{sy}$ equals one if medical marijuana is legal in state s and year y and $decriminalized_{sy}$ equals one if marijuana is decriminalized in state s and year y. I get an R-squared of .81. Regression results are presented in the appendix. I then impute missing values using the estimated coefficients. I use low-grade prices as the price individuals in my model face because it is unlikely that youth are purchasing and using medical-grade marijuana. Table 2 shows that marijuana prices vary over time and across states, but not in a systematic way like beer and cigarette prices.

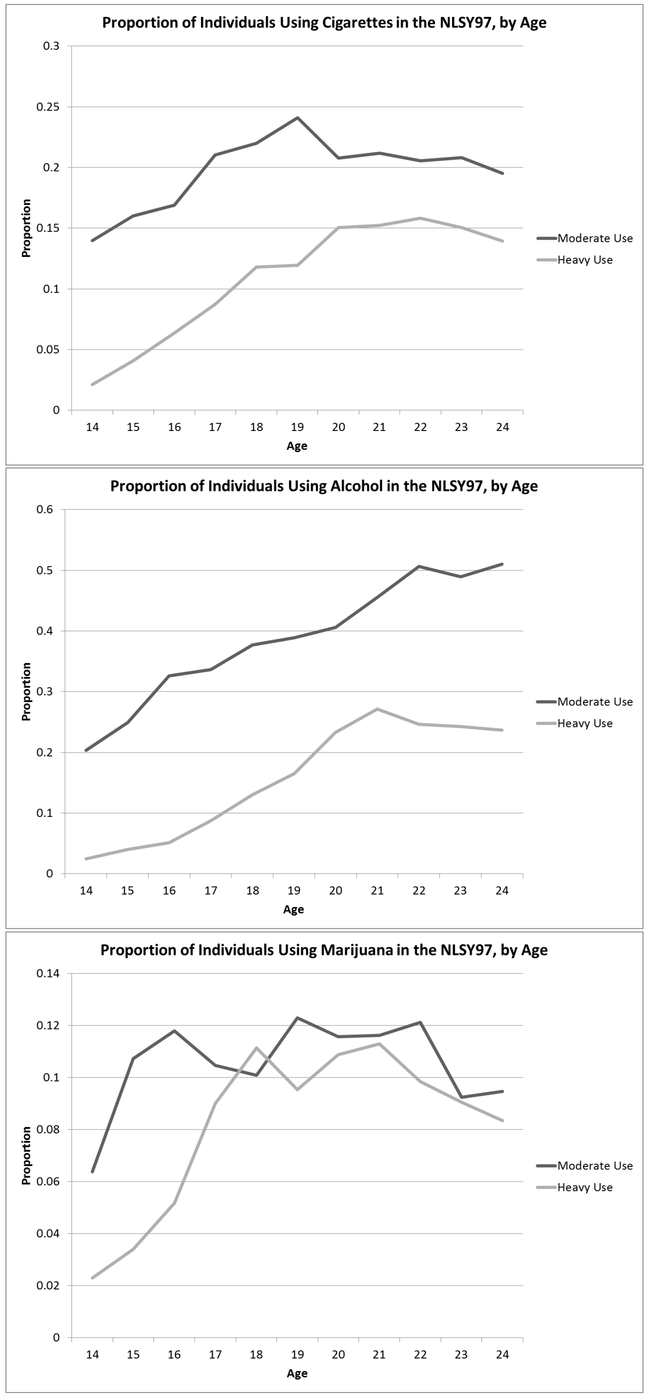
For the price of college, I use data collected from several years of the National Tuition and Fee Report published by the Washington Student Achievement Council.[14] The survey collects information on the average undergraduate tuition at over 200 state public institutions. I use the resident undergraduate tuition and required fees for 30 semester credit hours at the state's flagship university as the price of college.

4.3 Descriptive Statistics

This subsection discusses some key characteristics about substance use among individuals in my sample. Figure 2 shows the proportion of individuals using each substance by age. I classify individuals as using either none of a substance, a moderate amount, or a heavy amount. Details of how I classify use are presented in Figure 1. In general, sub-

[14] Andrew Barr provided me with this data.

Figure 2: Proportion Using Substances in the NLSY97, by Age



Data come from the NLSY97. The sample used is described in Table 1.

Table 2: Substance Price Summary Statistics, by Year

	Cigarette		Beer		Marijuana	
Year	Mean	Std. Dev.	Mean	Std Dev.	Mean	Std. Dev.
1997	2.64	(0.49)	4.65	(0.39)	123.62	(37.81)
1998	2.78	(0.57)	4.62	(0.37)	113.10	(47.35)
1999	2.97	(0.61)	4.63	(0.37)	134.62	(37.66)
2000	3.72	(0.62)	4.33	(0.39)	110.83	(30.21)
2001	3.81	(0.65)	4.27	(0.41)	124.13	(49.35)
2002	4.04	(0.66)	4.34	(0.44)	115.98	(44.38)
2003	4.38	(0.89)	4.34	(0.44)	123.54	(35.99)
2004	4.41	(0.91)	4.35	(0.44)	118.94	(49.46)
2005	4.39	(0.98)	4.23	(0.34)	108.66	(42.52)
2006	4.45	(1.01)	4.17	(0.37)	103.06	(40.80)
2007	4.41	(1.06)	4.24	(0.41)	124.05	(38.19)
2008	4.52	(1.09)	4.15	(0.42)	105.89	(30.43)

Data on cigarette prices come from volume 46 of The Tax Burden on Tobacco and are for one pack of cigarettes. A pack of cigarettes generally contains 20 cigarettes. Data on beer prices come from ACCRA's Cost of Living Index and are for a 6-pack of beer. Data on marijuana prices come from High Times Magazine and are for one ounce of Marijuana. I use the CPI to adjust all prices to 2000 dollars.

stance use increases as individuals get older. The proportion of individuals using a moderate amount of cigarettes increases until age 19, where 24 percent of my sample reports using a moderate amount of cigarettes. The proportion using a heavy amount of cigarettes increases to its peak of around 15 percent at age 20. The proportion drinking moderately increases until it flattens at age 22, where 50 percent of my sample reports drinking moderately. On the other hand, the proportion reporting drinking heavily hits its peak at age 21 at 27 percent. The percentage of individuals using moderate amounts of marijuana is around 12 percent for almost my entire sample, whereas those using heavy amounts increases drastically from 2.3 percent at age 14 to 10.9 percent at age 21.

I run several regressions to see how substance prices are associated with use at age 14. Table 3 presents the estimates and the marginal effects of an ordered probit of substance use on prices.[15] The dependent variable is the level of use at age 14: no use, moderate use, or heavy use. I assume that individuals have not used substances prior to age 14, so these regressions measure how prices affect the initiation of substance use. Marginal effects predict the probability of an individual *not using*. Table 3 shows that a one dollar increase

Table 3: Ordered Probit: Substance Use at Age 14

	Cigarette		Alcohol		Marijuana	
	Estimates	Marginal Effects	Estimates	Marginal Effects	Estimates	Marginal Effects
Cigarette Price	-0.187**	0.045**	0.167**	-0.049**	0.233**	-0.036**
	(0.09)	(0.02)	(0.08)	(0.02)	(0.10)	(0.02)
Beer Price	0.074	-0.018	-0.141	0.041	-0.234*	0.036*
	(0.10)	(0.02)	(0.09)	(0.03)	(0.12)	(0.02)
Marijuana Price	-0.001	0.000	-0.003***	0.001***	-0.002	0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Black	-0.499***	0.119***	-0.429***	0.126***	-0.063	0.010
	(0.12)	(0.03)	(0.11)	(0.03)	(0.13)	(0.02)
Hispanic	-0.190*	0.046*	-0.042	0.012	-0.097	0.015
	(0.11)	(0.03)	(0.10)	(0.03)	(0.13)	(0.02)
N	1,175		1,175		1,175	

Marginal effects predict the probability of an individual not using. Cigarette price reflects the price of a pack of cigarettes. A pack generally contains 20 cigarettes. Beer price reflects the price of a 6-pack of beer. Marijuana price reflect the price of an ounce of marijuana. A joint of marijuana contains around half of a gram.
* p<0.10, ** p<0.05, *** p<0.01

in the price of a pack of cigarettes is associated with a 4.5 percentage point increase in the probability that an individual *does not* use cigarettes at the age of 14. The price of beer is not statistically significantly associated with whether an individual drinks, but a dollar increase in the price of a pack of cigarettes is associated with a 4.9 percentage point decrease in the probability of *not drinking* at age 14. Additionally, a one dollar increase in the price of low-grade marijuana is associated with a 0.1 percentage point increase in the probability of *not drinking*. Lastly, I show that a dollar increase in cigarette prices is associated with a 3.6 percentage point decrease in the probability of *not using marijuana*;

[15] I do not present all of the marginal effects in this section. Please contact the author if you would like to see additional marginal effects.

a one dollar increase in the price of beer is associated with a 3.6 percentage point increase in the probability of *not using*. The price of marijuana is not statistically significantly associated with the individual's marijuana choice at age 14.

Next, I run several regressions to see how past substance use is associated with current substance use, work, and education outcomes at age 23. Results are presented in the appendix. None of the relationships I describe should be interpreted as causal, but they do describe key patterns in the data. The main concern is that the enjoyment that an individual gets from substance use is likely correlated with the educational and employment outcomes of interest for other reasons. The measures of past substance use that I use in these regressions are the total number of years the individual chose to consume any amount of a substance. For all substances, past substance use increases the probability of using that substance at age 23. However, past cigarette, alcohol, and marijuana use do not affect the probability of using the other substances. I find that past cigarette use decreases educational attainment at age 23, but past alcohol and marijuana use increase it. This counterintuitive result may arise because, for example, those who use marijuana and do not get arrested also have some sort of unobserved characteristic that is correlated with higher education. Past cigarette use decreases the probability of working while past marijuana use increases. Past cigarette and marijuana use do not statistically significantly affect the wages of those working at age 24; past alcohol use increases them.

5 Estimation Strategy

I estimate the parameters of my model using Bayesian Markov Chain Monte Carlo (MCMC) methods. Using classical estimation techniques to estimate my model is difficult for several reasons. First, it is important in my model that I allow for a substantial amount of time-persistent unobserved heterogeneity in order to identify the causal effects of state dependence on outcomes. It is generally burdensome to incorporate unobserved heterogeneity using classical methods because these terms must be integrated out in order to calculate the objective function, yet they generally do not have closed form solutions. MCMC estimation offers the convenience of data augmentation that allows me to avoid integration when evaluating my value function. Second, estimation with classical methods often require evaluating the value function for each sample observation and for each trial guess of the parameters. This can be very time consuming in problems with large choice sets and large state spaces. In particular, my model suffers from the curse of dimensionality due to the three continuous state variables pertaining to substance use stocks. Imai, Jain, and Ching (2009) (IJC) develop an estimator that approximates the value functions by using stored value functions from earlier iterations of the MCMC algorithm, making estimation feasible.[16] Lastly, my likelihood function is highly non-linear and probably not globally concave. This can make finding a global maximum difficult. The MCMC method is theoretically guaranteed to converge to the posterior distribution, which can be used to calculate the global maximum of the likelihood function.

5.1 Econometric Specification

Before discussing the estimation of my model in more detail, I specify the distributions of the random variables as well as the functional forms of the utility flow equations.

[16] Keane and Wolpin (1994) and Rust (1997) discuss ways to break the curse of dimensionality. The method developed in Rust (1997) does not apply to problems where the continuous state variable is deterministic, such as mine. Neither paper addresses the issue of allowing for substantial unobserved heterogeneity.

5.1.1 Distributional Assumptions

I assume that permanent worker-firm productivity match values and true randomness in wages are distributed as follows:

$$\begin{aligned} \psi_{ik} &\sim iidN\left(0, \sigma_{\psi}^{2}\right) \\ \nu_{itk}^{h} &\sim iidN\left(0, \sigma_{h}^{2}\right). \end{aligned}$$

Let $\psi_i = \{\psi_{ik}\}_{k=1}^{K_i}$ denote the vector of all job match offers an individual receives in his lifetime. Since an individual receives a new full-time and part-time job offer each period, the number of job matches an individual receives over his lifetime is $K_i = 2 \cdot T$.

I assume that the random shocks to the substance use decisions ϵ_{it}^{sub} are independent across individuals and time. I assume the random shocks to the human capital choices ϵ_{it}^{hc} are independent across individuals, time, and discrete choices. Specifically, $\epsilon_{it}^{sub} \sim iidN\left(0, \Sigma_{sub}\right)$ and $\epsilon_{it}^{hc} \sim iid\ EV$. As is true in all models with discrete choices, the parameters in my model are identified only up to a scale, so I set the variance of ϵ_{it}^{hc} equal to $\frac{\pi^2}{6}$.

I assume the parental transfer shocks are distributed log-normally and are independent across individuals, time, and discrete choices. I let the mean and standard deviation vary by the discrete hours and enrollment decisions; that is,

$$\ln\left(\tilde{w}_{it}\left(hc_{it}\right)\right) \sim iidN\left(\mu_{\tilde{w}}\left(hc_{it}\right), \sigma_{\tilde{w}}^{2}\left(hc_{it}\right)\right).$$

Lastly, I cannot separately identify an individual's time-persistent heterogeneity in preferences for nonemployment, part-time work, and full-time work. I set $\mu_i^{h=0}$ to zero, and I assume that the time-persistent individual heterogeneity in productivity μ_i^w and in preferences $\mu_i^N = \left(\mu_i^{sub}, \mu_i^P, \mu_i^F, \mu_i^s, \mu_i^{arrest}, \mu_i^{pass}\right)$

$$\mu_i = \begin{pmatrix} \mu_i^w \\ \mu_i^N \end{pmatrix} \sim N\left(0, \Sigma_\mu\right)$$

5.1.2 Utility Flow Equations

In this section, I parameterize the deterministic parts of the pecuniary and nonpecuniary utility flows. Table 4 summarizes the variables I include in each equation. Specific functional forms can be found in the appendix. Table 4 additionally clarifies the exclusion restrictions which I discussed in Section 3.4 and which help identify my model.

Table 4: Empirical Specification for Deterministic Parts of the Utility Function

	Pecuniary Utility	Nonpecuniary Utility		
Variables Included In Each Equation	Wage Equation	Working	School	Substance Use
Hours Part-Time	Yes*	Yes*	Yes	
Hours Full-Time			Yes	
Experience	Yes	Yes		
Experience Squared	Yes			
Tenure	Yes	Yes		
Worked Part-Time Last Period		Yes		
Worked Full-Time Last Period		Yes		
Education	Yes	Yes	Yes	
Enrolled Last Period			Yes	
Passed Last Period			Yes	
Substance Use	Yes	Yes	Yes	Yes
Substance Use Interacted with Other Use				Yes
Substance Use Interacted with Substance States				Yes
Substance Use Interacted with Age				Yes
Unemployment Rate	Yes	Yes		
Arrests	Yes	Yes	Yes	
Constant	Yes	Yes	Yes	Yes

* The wage and nonpecuniary utility are multiplied by the work part-time coefficient when individuals choose to work part-time.

The elements of the deterministic part of the wage equation $\theta_w\left(sub_{it}, s_{it}\right)$ are presented in the first column. Including experience and tenure in the wage equation captures how general and firm-specific human capital affect wages. Substance use history is included because it may limit productivity. Arrests are included so that I can separately identify the effect of substance use on wages from the effect of being arrested. Substance use is associated with more arrests; so, if I exclude arrests and arrests decrease productivity, then the effect of substance use on wages may be downward-biased. I include the unemployment rate because the unemployment rate may affect the distribution of wage offers.

The deterministic part of the nonpecuniary utility of employment, $\theta_N\left(sub_{it}, s_{it}\right)$, is presented in the second column. The deterministic part of the nonpecuniary utility from not working, $\theta_N^{h=0}\left(sub_{it}, s_{it}\right)$ is set equal to zero because the coefficients in the utility flow equation are identified only relative to a base choice. Current substance use is included because there may be a nonpecuniary cost to working if an individual is using a lot of substances.

The nonpecuniary cost function of attending school, $\kappa_s\left(h_{it}, sub_{it}, s_{it}\right)$, is presented in Column 3. I include whether the individual was enrolled last period in order to capture the cost of going to school after a period of not enrolling. An indicator of whether he passed if he was enrolled is also included; it is likely that an individual who did not pass last period has a higher nonpecuniary cost to enrolling this period. The cost of enrollment is allowed to vary by hours worked. Current substance use is included because an individual using large amounts of alcohol may have higher nonpecuniary costs of enrolling in school than an individual who is not drinking.

The nonpecuniary utility gained from using substances, $\alpha\left(sub_{it}, s_{it}\right)$ is presented in Column 4. I include the interaction between substance use and substance states to capture the addictiveness and gateway effects of substances. For example, a positive coefficient on the interaction of today's heavy cigarette use and past cigarette stock suggests that cigarettes are addictive and that there is a higher cost to quitting if an individual has used high amounts of the substance in the past. A positive coefficient on the the interaction of today's marijuana use and past cigarette use suggests that cigarette use is a gateway to marijuana use. Interactions between current substance uses are included to capture patterns such as individuals enjoying to smoke when they drink. These addiction, gateway, and complementarity effects are assumed to affect the utility of substance use but not the utility of employment or education choices conditional on substance use. Lastly, I include an indicator for whether or not an individual is using which captures any fixed costs associated with using a substance.

5.1.3 State Transition Probabilities

The probabilities of completing a grade of school and of getting arrested can be expressed as probit models with latent variables $G^*_{i,t+1}$ and $R^*_{i,t+1}$, respectively. That is,

$$
\begin{aligned}
G^*_{i,t+1} &= \eta_{school}\left(sub_{it}, s_{it}\right) + \mu_i^{pass} + N\left(0,1\right) \\
R^*_{i,t+1} &= \eta_{arrest}\left(sub_{it}, s_{it}\right) + \mu_i^{arrest} + N\left(0,1\right)
\end{aligned}
$$

where these probabilities depend on many of the state variables. I assume

$$\begin{aligned}\eta_{school}\left(sub_{it}, s_{it}\right) &= \eta_1^s + \eta_2^s Years\ of\ High\ School_{it} + \eta_3 Years\ of\ College_{it} + \\ &\quad \eta_4^s \mathbb{1}\left(G_{it} > 16\right) \eta_5^s enroll_{i,t-1} + \eta_6^s g_{i,t-1} enroll_{i,t-1} + \\ &\quad \eta_7^s \mathbb{1}\left(cigs_{it} = 2\right) + \eta_8^s \mathbb{1}\left(cigs_{it} = 3\right) + \\ &\quad \eta_9^s \mathbb{1}\left(alc_{it} = 2\right) + \eta_{10}^s \mathbb{1}\left(alc_{it} = 3\right) + \\ &\quad \eta_{11}^s \mathbb{1}\left(mj_{it} = 2\right) + \eta_{12}^s \mathbb{1}\left(mj_{it} = 3\right)\end{aligned} \tag{16}$$

and

$$\begin{aligned}\eta_{arrest}\left(c_{it}, d_{it}, s_{it}\right) &= \eta_1^a + \eta_2^a age_{it} + \eta_3^a age_{it}^2 + \eta_4^a R_{it} + \eta_5^a X_{it}^R + \\ &\quad \eta_6^a \mathbb{1}\left(cigs_{it} = 2\right) + \eta_7^a \mathbb{1}\left(cigs_{it} = 3\right) + \\ &\quad \eta_8^a \mathbb{1}\left(alc_{it} = 2\right) + \eta_9^a \mathbb{1}\left(alc_{it} = 3\right) + \\ &\quad \eta_{10}^a \mathbb{1}\left(mj_{it} = 2\right) + \eta_{11}^a \mathbb{1}\left(mj_{it} = 3\right).\end{aligned} \tag{17}$$

X_{it}^R are local law enforcement characteristics for individual i, which are not included elsewhere in the model. With this specification, a person completes a grade of school ($g_{it} = 1$) if $G_{i,t+1}^* > 0$ and a person gets arrested ($r_{it} = 1$) if $R_{i,t+1}^* > 0$. Therefore, conditional on $G_{i,t+1}^*$ and $R_{i,t+1}^*$ the transition of the state variables from period t to period $t + 1$ is deterministic.

5.2 Estimation Algorithm

As in most dynamic structural estimation algorithms, my problem can be divided into an outer loop and an inner loop. The outer loop estimates the parameters, while the inner loop calculates value functions that are used in the outer loop to calculate the objective function. I use Bayesian estimation in the outer loop to estimate the parameters of the model. Traditional approaches use backwards recursion in the inner loop to solve for the value functions. The IJC algorithm provides a way to approximate the value functions by using stored value functions from earlier iterations of the MCMC algorithm, making estimation more likely to be feasible. The rest of this subsection presents details of the outer and inner loops.

5.2.1 The Outer Loop

In Bayesian estimation, rather than estimating the parameters by optimizing an objective function, the researcher does so by simulating from the posterior distribution of the parameters. The posterior distribution is made up of the likelihood function and any prior beliefs the researcher has about the parameters. By Bayes' rule, the posterior distribution $\Pr(\theta|data) = \frac{\Pr(data|\theta)\Pr(\theta)}{\Pr(data)}$, where $\Pr(data|\theta)$ is the likelihood function, and $\Pr(\theta)$ are the prior beliefs. The approach is to make many draws of the parameters from the posterior distribution so that moments of the posterior can be calculated.

Usually, the posterior distribution does not have a convenient form from which to take draws. In this case, the MCMC method can be used to generate draws from the posterior distribution. To estimate the parameters of my model, I use Gibbs sampling in conjunction with the Metropolis-Hastings (M-H) algorithm and data augmentation.[17] Instead of taking draws from the multi-dimensional posterior of all the parameters, Gibbs sampling allows me to take draws of a subset of parameters called blocks, conditional on the values of the other parameters. The M-H algorithm is utilized in the case that the conditional posterior for a block of parameters does not take a simple form that can be easily drawn from.

In order to make calculation of the likelihood easier, I also use a data augmentation step within my Gibbs sampler. The usefulness of data augmentation was first discussed in Albert and Chib (1993). McCulloch and Rossi (1994) then showed how to use data augmentation within a Gibbs sampler to estimate a static multinomial probit model. Instead of simulating the posterior of the parameters, I simulate the joint posterior distribution of the parameters and latent variables, such as unobserved wage offers and time-persistent unobserved heterogeneity. In the data augmentation block, the latent variables are simulated conditional on the parameters and the data. Then, in the parameter block, the Gibbs sampler simulates the parameters conditional on the data and the latent variables. So, in this step the data has been augmented to include the latent variables, which makes drawing the parameter values easier by eliminating the need to perform high-dimensional integration. After repeating these steps, the resulting sequence of simulated parameters and latent variables is a Markov chain with the stationary distribution equal to the joint posterior distribution of the parameters and the latent variables.

[17] Chapters 9 and 12 in Train (2009) present a good introduction to Bayesian estimation, MCMC estimation, Gibbs sampling, and the M-H algorithm. Geman and Geman (1984), Gelfand and Smith (1990), Casella and George (1992), Smith and Roberts (1993), and Tierney (1994), among many others, discuss M-H and Gibbs sampling.

5.2.2 The Inner Loop

Within the blocks of the outer loop, I need to calculate the choice-specific value functions in order to calculate the likelihood function; this is the inner loop. As seen in Equation 14 (the choice-specific value function) to do this I need to be able to calculate the expected value function $EV\left(s_{i,t+1}|s_{it}, hc_{it}, sub_{it}\right)$. Following the IJC method, I approximate the expected value function by using saved information from earlier iterations of the MCMC estimation algorithm.[18] At the end of each MCMC iteration, I randomly pick one person and I calculate and store pseudo value functions for that person at each point in the discrete state space. To approximate the expected value function, I use a weighted average of the stored pseudo value functions, where an observation gets a higher weight if the current parameter values are close to the parameters that were used to calculate the stored value function.

To be more specific, at the end of MCMC iteration r, I have stored vectors of pseudo value functions $\left\{\left\{\tilde{V}^l_{i't}\left(s^d_t, s^{cl}_t|\theta^{*l}_{i'}\right)\right\}_{s^d_t \in S^d_t}\right\}^T_{t=1}$ for Q people, where i' is a randomly selected individual whose pseudo-value function was stored at the end of the lth iteration, $\theta^{*l}_{i'}$ are the parameters and latent variables from the lth iteration of the MCMC algorithm, and s^{cl}_t is a random draw of the three continuous state variables drawn from an uniform distribution over the support of the continuous state variables. The pseudo-expected value function is

$$\widetilde{EV}\left(s_{i,t+1}|s_{it}, hc_{it}, sub_{it}\right) = \sum_{l=r-Q}^{r-1} \tilde{V}^l_{i't}\left(s^d_t, s^{cl}_t|\theta^{*l}_{i'}\right) \cdot K_{H_\theta}\left(\theta^{*l}_{i'} - \theta^{cr}_i\right) \cdot K_{H_s}\left(s^{cl}_t - s^{cr}_t\right), \tag{18}$$

[18] See Ching et. al (2012) for a more detailed description of the method. Also, see Osborne (2011) and Zhou (2011) for applications of the algorithm.

where K_H is a Gaussian Kernel with bandwidth matrix H. This is simply a weighted average of the stored pseudo-value function where the weights are determined by how close the parameter variables, latent variables, and continuous state variables are to the current values.

It is important to note that random draws of the idiosyncratic shocks are used in the calculation of the stored pseudo value functions $\tilde{V}^{l}_{i't}\left(s^d_t, s^{cl}_t|\theta^{*l}_{i'}\right)$. However, these variables are not included in $\theta^{*l}_{i'}$, because these values are unknown to the agent at time period $T-1$. This method allows those variables to be integrated out. Lastly, this method allows me to use kernel-based local interpolation to overcome the curse of dimensionality caused by the three continuous state variables. As seen in Equation 18, I am weighing how close the continuous state variables are to previous values. Doing this allows me to calculate the pseudo value functions at only one randomly drawn vector of the continuous state space per parameter update.

Further details about my estimation strategy can be found in the appendix.

6 Results

In this section, I present estimates of the structural parameters of my model. These estimates were obtained from estimating my model on my sample of 642 white males, 282 black males, and 246 Hispanic males. I currently simplify the model in several ways.[19] First, I assume that the parameters that control how the addictive stocks evolve $(\beta^c, \beta^a, \beta^m) = 0.5$. I assume that the probability that a part-time job increases human capital by one unit π^H equals one. Additionally, I assume the random shocks to the substance use decisions

[19]Relaxing the simplifications is not difficult. Most of these simplifications were made in order to speed up estimation.

are $\epsilon_{it}^{sub} \sim iidN(0,1)$, rather than estimating Σ_{sub}. Lastly, I assume that the minimum level of consumption $C_{min} = \$11,000$. If an individual were to make a choice that would put their consumption below C_{min}, then I provide them with enough income so that that $CONS_{it} = C_{min}$. I ran the MCMC estimation algorithm for approximately 50,000 iterations. I used the 40,000th to the 50,000th iterations to derive the posterior means and standard errors, which are presented below. The standard errors of the parameters are simply the standard deviations of the posterior distribution. Results are presented in Tables 5 through 13. There are too many parameters to discuss each one individually, so instead I focus my discussion on key parameter estimates and what the estimates reveal about the substance use and human capital accumulation decision process.

6.1 Nonpecuniary Utility Flows

The estimated parameters of the nonpecuniary utility flows are presented in Tables 5 through 7. These parameter estimates represent the effect of each variable on the nonpecuniary utility of enrolling in school, working part-time, working full-time, and using substances. All estimates are relative to the base choice of not enrolling, not working, and not using substances.

I first present the estimates of the nonpecuniary utility of enrolling in school in Table 5. The results show that an extra year of high school education increases the utility of enrolling in school by 2.690 log-yearly consumption units for white males (3.644 for Hispanic males and 1.686 for black males). On the other hand, an extra year of college education decreases the utility of enrolling in school by 1.939 log-yearly consumption units for white males, 1,564 for Hispanic males, and 1.031 for black males. Working part-time raises the nonpecuniary utility of attending school for white and black males and so is a complement to enrollment, whereas working full-time is a substitute. Having been arrested decreases the utility of enrolling in school. Consuming a heavy amount of cigarettes has a negative sign and so is a substitute for enrolling in school for all males. For these individuals, heavy cigarette use is equivalent to decreasing yearly consumption by 51.6 percent for white males, 66.24 percent for Hispanic males, and 55.5 percent for black males. Using a heavy amount of alcohol is a complement to school enrollment for Hispanic males, but not for white or black males. Consuming heavy amounts of marijuana increases the utility of enrolling for white and black males, but not for Hispanic males.

Table 5: Non-Pecuniary Utility of Attending School Coefficients

	White		Hispanic		Black	
	mean	se	mean	se	mean	se
Years of High School Completed	2.690*	(0.10)	3.644*	(0.08)	1.686*	(0.14)
Years of College Completed	-1.939*	(0.06)	-1.564*	(0.03)	-1.031*	(0.02)
Enrolled Last Period	7.653*	(0.18)	6.153*	(0.32)	9.919*	(0.27)
Passed Last Period	10.932*	(0.53)	6.749*	(0.40)	0.588*	(0.25)
Working Part-time	1.164*	(0.28)	-0.075	(0.18)	-0.895*	(0.27)
Working Full-time	-0.843*	(0.27)	-1.279*	(0.16)	-2.278*	(0.18)
Arrests	-0.500*	(0.05)	-0.492*	(0.02)	-0.350*	(0.05)
Moderate Cigarettes	-0.047	(0.06)	-0.483*	(0.15)	-0.397*	(0.13)
Heavy Cigarettes	-0.725*	(0.08)	-1.086*	(0.24)	-0.812*	(0.19)
Moderate Alcohol	0.290*	(0.05)	0.567*	(0.11)	0.249*	(0.11)
Heavy Alcohol	0.097	(0.09)	1.203*	(0.13)	-0.182	(0.12)
Moderate Marijuana	0.352*	(0.09)	-0.196	(0.14)	-0.088	(0.16)
Heavy Marijuana	0.423*	(0.09)	-0.864*	(0.19)	0.469*	(0.19)
Constant	-13.898*	(0.28)	-16.669*	(0.41)	-16.095*	(0.30)

* $p<0.05$

I present the estimates of the nonpecuniary utility of working in Table 6. Individuals get less nonpecuniary utility from working part-time than they do from working full-time (because the coefficient is less than one). An extra year of experience or tenure increases the utility of working. An extra year of education decreases the utility of working for white males, but increases it for black males. Individuals who worked in the previous period have a higher utility of working in the current period. The coefficients on moderate alcohol use are all positive, suggesting that moderate alcohol use and working are complements; a possible explanation for this is that individuals value being able to have a drink with co-workers after work. Heavy cigarette and alcohol use increase the nonpecuniary utility of working for white males. Marijuana use decreases the utility of working for Hispanic males.

Table 6: Non-Pecuniary Utility of Working Equation Coefficients

	White		Hispanic		Black	
	mean	se	mean	se	mean	se
Experience	0.310*	(0.01)	0.311*	(0.03)	0.111*	(0.05)
Tenure	0.217*	(0.01)	0.235*	(0.04)	0.099*	(0.05)
Education	-0.114*	(0.02)	-0.025	(0.06)	0.388*	(0.02)
Arrests	-0.210*	(0.03)	-0.462*	(0.05)	-0.220*	(0.04)
Worked Part-time Last Period	4.948*	(0.30)	4.099*	(0.22)	5.698*	(0.42)
Worked Full-time Last Period	6.686*	(0.17)	5.087*	(0.35)	6.891*	(0.21)
Unemployment Rate	-0.543*	(0.01)	-0.126*	(0.06)	-0.860*	(0.03)
Moderate Cigarettes	-0.192	(0.11)	0.154	(0.11)	-0.067	(0.14)
Heavy Cigarettes	0.221*	(0.08)	-0.146	(0.30)	0.334	(0.19)
Moderate Alcohol	0.499*	(0.13)	0.270*	(0.12)	0.250*	(0.12)
Heavy Alcohol	0.373*	(0.12)	0.241	(0.16)	-0.031	(0.19)
Moderate Marijuana	0.138	(0.10)	-0.515*	(0.12)	-0.010	(0.16)
Heavy Marijuana	0.197	(0.14)	-0.507*	(0.16)	-0.209	(0.16)
Constant	-4.284*	(0.15)	-5.183*	(0.33)	-7.081*	(0.23)
Work Part-Time***	0.473*	(0.03)	0.305*	(0.07)	0.257*	(0.06)

*** The utility of working is multiplied by the work part-time coefficient when individuals choose to work part-time.
* $p<0.05$

Lastly, I present the estimates of the nonpecuniary utility associated with using substances in Table 7. The coefficients on the interaction of substance use and the addictive stock being equal to zero show there is a large nonpecuniary cost of starting to use a substance for the first time. In particular, there is a significant and large nonpecuniary cost to using a heavy amount of a substance if you have never used that substance before. For example, using a heavy amount of alcohol, if a person has never used alcohol before, is equivalent to a 74.3 to 83.2 percent decrease in yearly consumption compared to someone who has used alcohol before. Increasing the addictive stock of a substance increases the utility of using a heavy amount of that substance, suggesting that all substances are addictive. The positive coefficient on the interaction of current cigarette and marijuana use suggests that cigarettes and marijuana are complements. However, past cigarette use decreases the utility of consuming marijuana, suggesting that marijuana and cigarettes may be substitutes. Past alcohol use increases the utility of using moderate amounts of marijuana, suggesting that alcohol is a gateway drug and that alcohol and marijuana are complements. I additionally find evidence supporting the reverse gateway theory, because I find that past marijuana use increases the utility of using any level of cigarettes.

6.2 Wage Equation

Estimates of the log-wage equation parameters are presented in Table 8. White, Hispanic, and black males working part-time earn hourly wages 4.7, 5.0, and 2.7 percent lower, respectively, than individuals working full-time. Each extra year of high school education increases the wages of Hispanic and black males by 7.6 and 5.6 percent, respectively, but has no statistically significant effect on white workers. On the other hand, each extra year of college education increases the wages of white males by 4.1 percent. However, there is no statistically significant effect of college education on black and Hispanic males. This is likely due to the fact that I only observe individuals until they are 24, when many males have only recently completed college. The benefits of more education may not have had enough time to show up as higher wages. Each additional arrest decreases the wages of black males by 6.1 percent, but has no significant effect on white or Hispanic males. Cigarette and alcohol use decrease the wages of white and Hispanic males; marijuana use does not affect wages.

Table 7: Non-Pecuniary Utility of Consuming Drugs Coefficients

	White		Hispanic		Black	
	mean	se	mean	se	mean	se
Moderate Cigarette Use						
Constant	-1.864*	(0.06)	-1.087*	(0.12)	-1.472*	(0.08)
Cigarette Addictive Stock=0	-1.152*	(0.09)	-1.116*	(0.11)	-1.029*	(0.12)
Cigarette Addictive Stock	0.381*	(0.02)	0.463*	(0.01)	0.414*	(0.01)
Alcohol Addictive Stock	-0.099*	(0.01)	0.026	(0.02)	-0.344*	(0.01)
Marijuana Addictive Stock	0.291*	(0.01)	0.131*	(0.02)	0.329*	(0.01)
Age	-0.088*	(0.01)	-0.207*	(0.02)	-0.062*	(0.01)
Heavy Cigarette Use						
Constant	-1.955*	(0.05)	-2.311*	(0.24)	-4.265*	(0.13)
Cigarette Addictive Stock=0	-1.953*	(0.07)	-2.309*	(0.30)	-1.920*	(0.28)
Cigarette Addictive Stock	1.317*	(0.01)	1.221*	(0.01)	1.141*	(0.03)
Alcohol Addictive Stock	-0.740*	(0.02)	-0.732*	(0.01)	-0.672*	(0.02)
Marijuana Addictive Stock	0.308*	(0.01)	0.459*	(0.01)	0.378*	(0.02)
Age	-0.133*	(0.01)	-0.259*	(0.02)	0.065*	(0.02)
Moderate Alcohol Use						
Constant	-1.861*	(0.05)	-2.219*	(0.22)	-2.338*	(0.13)
Alcohol Addictive Stock=0	-0.707*	(0.05)	-0.585*	(0.10)	-1.303*	(0.08)
Cigarette Addictive Stock	-0.125*	(0.02)	0.173*	(0.02)	-0.222*	(0.03)
Alcohol Addictive Stock	0.332*	(0.02)	0.443*	(0.02)	0.124*	(0.01)
Marijuana Addictive Stock	-0.111*	(0.01)	0.041*	(0.02)	-0.050*	(0.01)
Age	0.142*	(0.01)	0.147*	(0.03)	0.181*	(0.02)
Heavy Alcohol Use						
Constant	-3.047*	(0.05)	-4.161*	(0.28)	-3.519*	(0.19)
Alcohol Addictive Stock=0	-1.514*	(0.08)	-1.358*	(0.10)	-1.786*	(0.19)
Cigarette Addictive Stock	-0.147*	(0.01)	-0.108*	(0.01)	-0.191*	(0.01)
Alcohol Addictive Stock	0.879*	(0.02)	0.909*	(0.02)	0.851*	(0.01)
Marijuana Addictive Stock	0.165*	(0.01)	0.126*	(0.01)	0.178*	(0.01)
Age	0.034*	(0.01)	0.094*	(0.01)	0.021	(0.01)
Moderate Marijuana Use						
Constant	-3.361*	(0.06)	-2.331*	(0.15)	-2.130*	(0.14)
Marijuana Addictive Stock=0	-0.835*	(0.06)	-0.975*	(0.22)	-1.442*	(0.10)
Cigarette Addictive Stock	-0.134*	(0.02)	-0.094*	(0.02)	-0.310*	(0.01)
Alcohol Addictive Stock	0.277*	(0.01)	0.127*	(0.01)	0.154*	(0.03)
Marijuana Addictive Stock	0.029	(0.02)	-0.149*	(0.01)	-0.189*	(0.01)
Age	-0.200*	(0.01)	-0.183*	(0.01)	-0.145*	(0.04)
Heavy Marijuana Use						
Constant	-1.517*	(0.05)	-3.667*	(0.09)	-3.982*	(0.12)
Marijuana Addictive Stock=0	-3.531*	(0.12)	-0.817*	(0.11)	-0.887*	(0.18)
Cigarette Addictive Stock	-0.384*	(0.01)	-0.477*	(0.01)	-0.271*	(0.03)
Alcohol Addictive Stock	-0.225*	(0.01)	-0.092*	(0.02)	-0.086*	(0.04)
Marijuana Addictive Stock	1.133*	(0.01)	1.139*	(0.02)	1.087*	(0.02)
Age	-0.187*	(0.01)	-0.085*	(0.01)	-0.063*	(0.01)
Using Cigarettes and Alcohol	1.243*	(0.05)	1.553*	(0.08)	1.263*	(0.10)
Using Cigarettes and Marijuana	1.722*	(0.05)	1.537*	(0.09)	1.601*	(0.09)
Using Alcohol and Marijuana	1.892*	(0.09)	1.770*	(0.09)	1.559*	(0.11)

* $p<0.05$. Age is measures as true age minus 13.

Table 8: Log-Wage Equation Coefficients

	White		Hispanic		Black	
	mean	se	mean	se	mean	se
Years of High School Completed	0.019	(0.01)	0.076*	(0.01)	0.056*	(0.02)
Years of College Completed	0.041*	(0.01)	0.036	(0.03)	0.022	(0.03)
Experience	0.106*	(0.01)	0.114*	(0.02)	-0.029*	(0.01)
Experience Squared	-0.004*	(0.00)	-0.010*	(0.00)	0.015*	(0.00)
Tenure	0.010	(0.02)	0.004	(0.02)	0.061*	(0.03)
Unemployment Rate	-0.004	(0.01)	-0.067*	(0.02)	0.008	(0.02)
Arrests	-0.016	(0.02)	-0.061	(0.03)	-0.034*	(0.01)
Moderate Cigarettes	-0.140*	(0.03)	-0.124*	(0.05)	-0.086	(0.05)
Heavy Cigarettes	-0.184*	(0.03)	-0.167*	(0.08)	-0.118	(0.08)
Moderate Alcohol	-0.165*	(0.03)	-0.154*	(0.04)	-0.051	(0.05)
Heavy Alcohol	-0.178*	(0.03)	-0.150*	(0.07)	-0.073	(0.07)
Moderate Marijuana	-0.041	(0.03)	-0.042	(0.06)	-0.054	(0.07)
Heavy Marijuana	-0.024	(0.03)	-0.003	(0.08)	0.025	(0.07)
Constant	2.127*	(0.05)	2.293*	(0.09)	1.866*	(0.16)
Work Part-Time***	0.953*	(0.01)	0.950*	(0.03)	0.973*	(0.03)
Error Standard Deviations						
True Randomness in Wages	0.379*	(0.00)	0.381*	(0.01)	0.363*	(0.01)
Firm Match Value	0.018*	(0.00)	0.004*	(0.00)	0.016*	(0.01)

*** The wage is multiplied by the work part-time coefficient when individuals choose to work part-time.
* $p<0.05$

6.3 State Transition Probabilities

In this subsection, I present estimates of the parameters that determine how the state variables evolve. These estimates can be found in Tables 9 and 10. First, I present the coefficients of the probability that an individual passes a grade in Table 9. An extra year of high school decreases the probability of passing for white males, but has no effect on the probability of passing for Hispanic and black males. An extra year of college decreases the probability of passing for all males. Individuals who enrolled and passed in the previous period have a higher probability of passing. Overall, substance use has very little effect on the probability of passing. Heavy alcohol and moderate marijuana use decrease the probability of passing for white males whereas heavy marijuana decreases the probability of passing for black males.

Next, I present the probability that an individual gets arrested in Table 10. For white males, the probability of arrest increases until they are 21, then decreases. As Hispanic and black males get older, the probability of arrest decreases. Having been previously arrested increases the probability of arrest. Heavy cigarette and alcohol use increase the probability of arrest for all males. Heavy marijuana use increases the probability of arrest for white and black males, but has no statistically significant effect on Hispanic males.

6.4 Heterogeneity in Preferences and Skills

Lastly, I present the variance-covariance matrix of the time-persistent unobserved heterogeneity in preferences and skills in Tables 11, 12, and 13. For white males, I find that individuals with higher than average preferences for heavy cigarette use also tend to have higher than average preferences for heavy alcohol and heavy marijuana use, lower preferences for enrolling in school, higher chances of getting arrested, and lower chances of pass-

Table 9: Probability of Passing if Enrolled Coefficients

	White		Hispanic		Black	
	mean	se	mean	se	mean	se
Years of High School Completed	-0.042*	(0.00)	-0.005	(0.00)	-0.006	(0.02)
Years of College Completed	-0.189*	(0.00)	-0.139*	(0.00)	-0.145*	(0.00)
Enrolled in Graduate School	-0.022	(0.03)	0.369*	(0.02)	0.697*	(0.07)
Enrolled Last Period	0.517*	(0.01)	0.405*	(0.02)	-0.013	(0.05)
Passed Last Period	1.560*	(0.02)	1.352*	(0.02)	1.296*	(0.04)
Moderate Cigarettes	0.002	(0.01)	-0.002	(0.01)	-0.016	(0.01)
Heavy Cigarettes	0.008	(0.01)	-0.017	(0.01)	0.040	(0.02)
Moderate Alcohol	0.000	(0.01)	0.001	(0.01)	0.005	(0.01)
Heavy Alcohol	-0.052*	(0.01)	-0.015	(0.02)	0.014	(0.01)
Moderate Marijuana	-0.020*	(0.01)	-0.005	(0.01)	-0.020	(0.01)
Heavy Marijuana	-0.009	(0.01)	-0.002	(0.01)	-0.119*	(0.02)
Constant	-1.101*	(0.01)	-1.003*	(0.02)	-0.376*	(0.03)

* p<0.05

Table 10: Probability of Arrest Coefficients

	White		Hispanic		Black	
	mean	se	mean	se	mean	se
Age	0.152*	(0.03)	-0.189*	(0.03)	-0.084*	(0.02)
Age Squared	-0.022*	(0.00)	0.005*	(0.00)	-0.001	(0.00)
Arrests	0.326*	(0.02)	0.489*	(0.01)	0.373*	(0.04)
Police Per Person	4.974	(11.01)	16.270	(13.03)	-36.626*	(18.24)
Moderate Cigarettes	0.439*	(0.08)	0.194	(0.13)	0.063	(0.10)
Heavy Cigarettes	0.587*	(0.07)	0.414*	(0.17)	0.517*	(0.14)
Moderate Alcohol	0.115	(0.07)	0.235*	(0.10)	0.149*	(0.06)
Heavy Alcohol	0.274*	(0.09)	0.505*	(0.16)	0.282*	(0.11)
Moderate Marijuana	0.258*	(0.07)	0.048	(0.14)	0.286*	(0.09)
Heavy Marijuana	0.230*	(0.08)	0.084	(0.24)	0.527*	(0.07)
Constant	-2.119*	(0.08)	-1.225*	(0.10)	-1.239*	(0.07)

* p<0.05. Age is measures as true age minus 13.

ing a grade. White males who have higher preferences for heavy alcohol have higher preferences for heavy marijuana, but lower preferences for moderate marijuana; have stronger preferences for working; have higher probability of arrest; and have lower probability of passing. Having a high preference for heavy marijuana use is positively correlated with a high preference for working, getting arrested, and is negatively correlated with passing.

For Hispanic males, having a high preference for heavy cigarette use is correlated with a high preference for heavy alcohol and marijuana use, working full-time, and a higher probability of passing. Individuals with strong preferences for heavy alcohol use have higher preferences for working, lower preferences for enrolling in school, a higher probability of arrest and a lower probability of passing a grade. Those with higher preferences for using heavy marijuana have higher preferences for working full-time, have higher probability of getting arrested, and a lower probability of passing.

Black males with higher than average preferences for heavy cigarettes are likely to have higher preferences for alcohol use, lower preferences for marijuana use, higher preferences for working, higher probability of passing, higher wages, and higher probability of arrest. Having a higher preference for heavy alcohol use is positively correlated with preferences for marijuana use and working full-time and is negatively correlated with preferences for enrolling in school. Individuals with higher preferences for heavy alcohol use also have a higher probability of passing a grade and higher earnings potential. Black males with higher preferences for using heavy marijuana have higher preferences for working full-time and enrolling in school, but a lower probability of passing.

Table 11: Variance-Covariance of Unobserved Heterogeneity for White Males

	Moderate Cigarettes	Heavy Cigarettes	Moderate Alcohol	Heavy Alcohol	Moderate Marijuana	Heavy Marijuana	Work Part-Time	Work Full-Time	School	Wage	Arrest	Pass
Moderate Cigarettes	1.098*											
	(0.12)											
Heavy Cigarettes	0.201*	0.829*										
	(0.06)	(0.05)										
Moderate Alcohol	-0.004	0.101*	0.567*									
	(0.04)	(0.04)	(0.04)									
Heavy Alcohol	0.351*	0.537*	-0.036	0.624*								
	(0.07)	(0.04)	(0.03)	(0.04)								
Moderate Marijuana	0.130*	0.051	-0.124*	-0.145*	0.511*							
	(0.05)	(0.04)	(0.04)	(0.03)	(0.04)							
Heavy Marijuana	0.101*	0.170*	0.037*	0.213*	-0.141*	0.164*						
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)						
Work Part-time	0.049	0.318*	0.299*	0.319*	-0.311*	0.128*	0.740*					
	(0.08)	(0.05)	(0.03)	(0.04)	(0.04)	(0.02)	(0.07)					
Work Full-time	0.285*	0.307*	0.196*	0.403*	-0.419*	0.233*	0.450*	0.636*				
	(0.09)	(0.04)	(0.04)	(0.05)	(0.04)	(0.02)	(0.04)	(0.05)				
School	-0.429*	-0.116*	-0.132*	-0.134*	-0.088*	-0.005	-0.133*	-0.106*	0.242*			
	(0.05)	(0.04)	(0.02)	(0.03)	(0.02)	(0.01)	(0.04)	(0.04)	(0.03)			
Wage	0.126*	0.144*	0.062*	0.097*	0.016*	0.031*	0.053*	0.081*	-0.064*	0.043*		
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)		
Arrest	0.064*	0.057*	0.024*	0.081*	0.027*	0.024*	0.078*	0.017	-0.056*	0.006*	0.050*	
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	
Pass	0.020	-0.015*	0.000	-0.030*	-0.025*	-0.015*	0.012	0.034*	-0.005	-0.001	-0.014*	0.023*
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)

* $p<0.05$

Table 12: Variance-Covariance of Unobserved Heterogeneity for Hispanic Males

	Moderate Cigarettes	Heavy Cigarettes	Moderate Alcohol	Heavy Alcohol	Moderate Marijuana	Heavy Marijuana	Work Part-Time	Work Full-Time	School	Wage	Arrest	Pass
Moderate Cigarettes	0.736* (0.11)											
Heavy Cigarettes	0.071 (0.09)	1.686* (0.19)										
Moderate Alcohol	-0.088 (0.06)	0.237* (0.09)	0.355* (0.06)									
Heavy Alcohol	0.432* (0.06)	0.343* (0.09)	-0.183* (0.06)	0.540* (0.06)								
Moderate Marijuana	0.032 (0.04)	0.494* (0.08)	0.006 (0.04)	0.059* (0.03)	0.247* (0.03)							
Heavy Marijuana	0.147* (0.04)	0.140* (0.06)	-0.007 (0.02)	0.243* (0.04)	-0.020 (0.02)	0.207* (0.03)						
Work Part-time	0.385* (0.07)	0.044 (0.10)	0.196* (0.04)	0.029 (0.06)	-0.064 (0.04)	-0.075 (0.04)	0.668* (0.09)					
Work Full-time	0.236* (0.04)	0.187* (0.06)	0.130* (0.03)	0.173* (0.04)	-0.061* (0.03)	0.126* (0.03)	0.257* (0.04)	0.254* (0.03)				
School	-0.322* (0.05)	0.067 (0.05)	-0.020 (0.03)	-0.098* (0.03)	0.022 (0.02)	0.002 (0.02)	-0.295* (0.04)	-0.118* (0.02)	0.190* (0.03)			
Wage	0.057* (0.01)	0.172* (0.02)	0.051* (0.01)	0.065* (0.02)	0.029* (0.01)	0.041* (0.01)	0.048* (0.01)	0.068* (0.01)	-0.021* (0.01)	0.031* (0.00)		
Arrest	0.228* (0.03)	0.021 (0.03)	-0.036* (0.02)	0.128* (0.02)	0.017 (0.01)	0.037* (0.01)	0.122* (0.02)	0.061* (0.02)	-0.099* (0.01)	0.013* (0.00)	0.075* (0.01)	
Pass	0.009 (0.02)	0.095* (0.03)	0.078* (0.02)	-0.080* (0.02)	0.037* (0.01)	-0.075* (0.01)	0.141* (0.03)	0.018 (0.01)	-0.041* (0.01)	0.008 (0.01)	0.007 (0.01)	0.064* (0.01)

* p<0.05

Table 13: Variance-Covariance of Unobserved Heterogeneity for Black Males

	Moderate Cigarettes	Heavy Cigarettes	Moderate Alcohol	Heavy Alcohol	Moderate Marijuana	Heavy Marijuana	Work Part-Time	Work Full-Time	School	Wage	Arrest	Pass
Moderate Cigarettes	0.951* (0.11)											
Heavy Cigarettes	-0.119 (0.08)	0.795* (0.09)										
Moderate Alcohol	-0.069 (0.05)	0.148* (0.04)	0.394* (0.05)									
Heavy Alcohol	0.314* (0.05)	0.278* (0.05)	-0.111* (0.04)	0.525* (0.06)								
Moderate Marijuana	0.055 (0.04)	-0.022 (0.06)	-0.175* (0.03)	0.036 (0.04)	0.358* (0.04)							
Heavy Marijuana	0.021 (0.04)	-0.138* (0.04)	-0.049 (0.03)	0.104* (0.03)	-0.065* (0.02)	0.227* (0.03)						
Work Part-time	-0.218 (0.12)	0.605* (0.09)	0.358* (0.06)	-0.069 (0.08)	-0.314* (0.06)	-0.238* (0.05)	1.319* (0.16)					
Work Full-time	0.023 (0.04)	0.200* (0.06)	0.172* (0.04)	0.084* (0.04)	-0.333* (0.04)	0.087* (0.03)	0.414* (0.07)	0.404* (0.05)				
School	-0.297* (0.05)	-0.062 (0.03)	-0.075* (0.02)	-0.065* (0.03)	-0.048* (0.02)	0.094* (0.02)	-0.173* (0.04)	0.010 (0.02)	0.191* (0.03)			
Wage	0.056* (0.02)	0.058* (0.02)	0.037* (0.01)	0.086* (0.01)	0.024* (0.01)	0.024* (0.01)	-0.079* (0.03)	-0.003 (0.01)	-0.012 (0.01)	0.047* (0.01)		
Arrest	0.055* (0.03)	0.064* (0.02)	-0.016 (0.02)	0.050* (0.02)	-0.008 (0.02)	-0.001 (0.02)	0.187* (0.04)	0.056* (0.02)	-0.057* (0.01)	-0.028* (0.01)	0.089* (0.01)	
Pass	-0.052 (0.04)	0.142* (0.03)	0.095* (0.02)	-0.137* (0.03)	-0.110* (0.02)	-0.121* (0.02)	0.418* (0.06)	0.126* (0.03)	-0.051* (0.02)	-0.055* (0.01)	0.054* (0.02)	0.195* (0.02)

* p<0.05

7 Model Fit

This section examines how well the structural model fits the data. To do this, I simulate school enrollment, employment, and substance use choices using the estimated parameters of my structural model. I compare these simulated choices with the choices observed in the data. Additionally, I compare the simulated probability of passing a grade and getting arrested with the proportion observed in the data. Lastly, I compare simulated log-wages generated by the structural model with those observed in the data. I present these comparisons in Figures 3 through 5.

Simulated data from the structural model differ from the actual data, although the model fits the overall patterns of enrollment, working, and substance choices quite well. I am able to capture the overall pattern of school enrollment for all males well. I underpredict enrollment of white males during ages 18 and 19. That is, I am underpredicting the number of individuals enrolling in college immediately after graduating from high school. I also underpredict enrollment of black males from age 14 to 15 and overpredict from age 18 to 20, but I fit the overall decreasing pattern of enrollment quite well. I am able to capture the pattern of passing rates for Hispanics and blacks, but underpredict the passing rate for white males across all ages.

I overpredict the proportion of white males working part-time and underpredict the proportion of working full-time. However, I am capturing the downward trend of working part-time and the upward trend of working full-time nicely. The simulated log-wages for white males are quite similar to the observed log-wages, through age 22, where I underpredict wages. Similarly, the simulated log-wages for Hispanic males match the data well until age 23. I also match the overall pattern of arrests well. Although, I overpredict the proportion of white males being arrested from ages 16-23. The coefficients on age and age squared may not be separately identified, causing me to overpredict the probability of arrest.

Lastly, I discuss the fit of substance use decisions. Overall, I fit the substance use choices well. For black males, I overpredict the proportion of individuals using heavy cigarettes after age 22. For white and Hispanic males, I overpredict the proportion using heavy alcohol after age 22. In all three situations, I am unable to capture the observed decrease in use after age 21. In my current model, the time-trend for use is linear in age. Relaxing this assumption by making it quadratic or allowing it to differ after age 21 may fix this problem.

8 Counterfactual Simulations

One appeal to estimating a structural model is the ability to run counterfactual simulations with the estimated model. In this section, I present counterfactual simulations that help explore the results better as well as simulations meant to provide policy recommendations. In the first subsection, I present a simulation that highlights the role that unobserved preferences play in the substance use decisions of individuals. In Subsection 8.2, I present simulations that make clear the interdependencies of cigarette, alcohol, and marijuana use. In the third subsection I present two policy experiments aimed at improving the outcomes of black and Hispanic males by lowering their probability of arrest. In the first experiment, I simulate what would happen if black and Hispanic males faced the same probability of arrest as white males. In the second, I simulate what would happen to black males if marijuana use was legal. It is important to note the simulations I am presenting below are partial equilibrium simulations. That is, they ignore any changes to the equilibrium that

Figure 3: Comparison of Simulated and Actual Choices for White Males

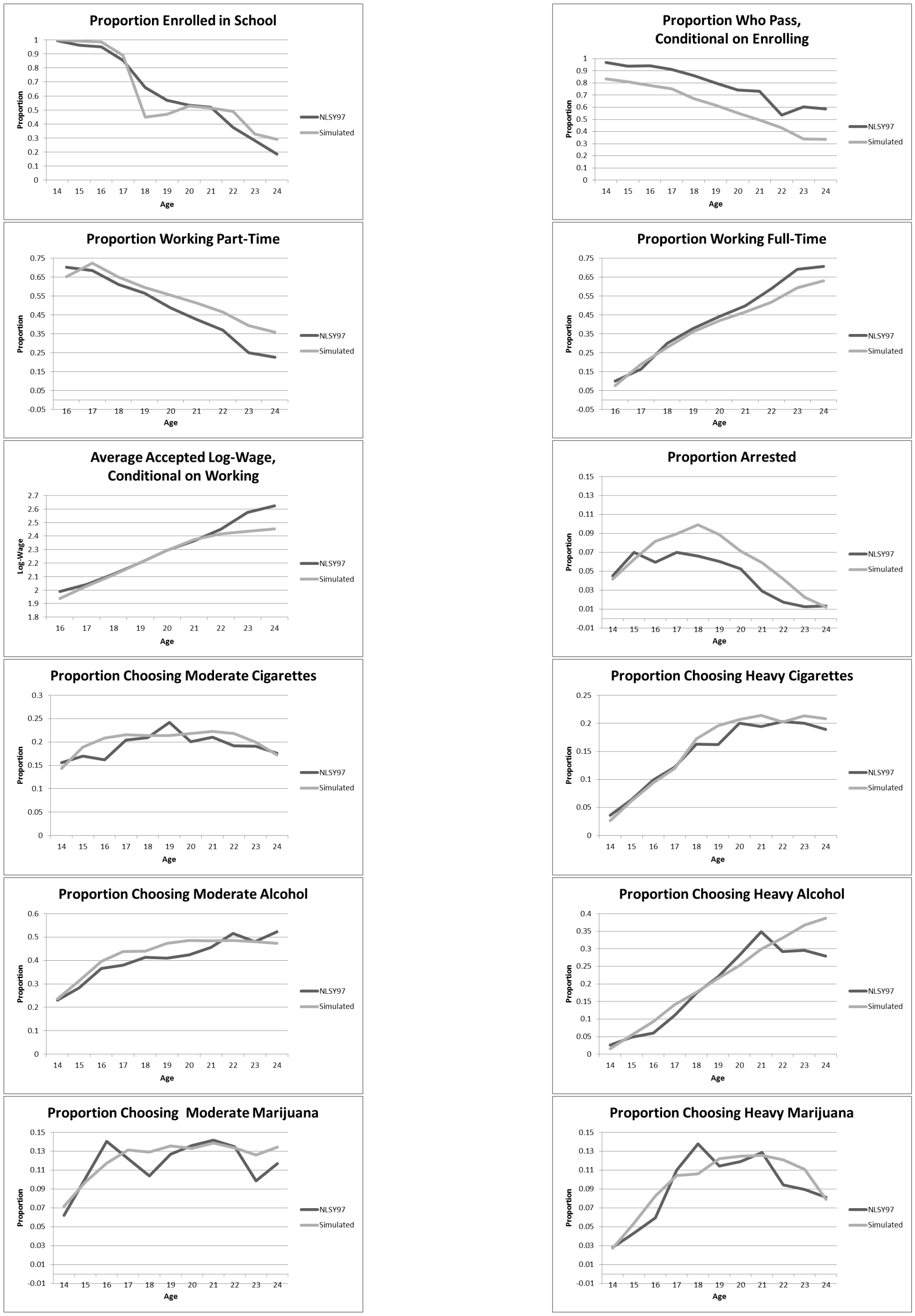


Figure 4: Comparison of Simulated and Actual Choices for Hispanic Males

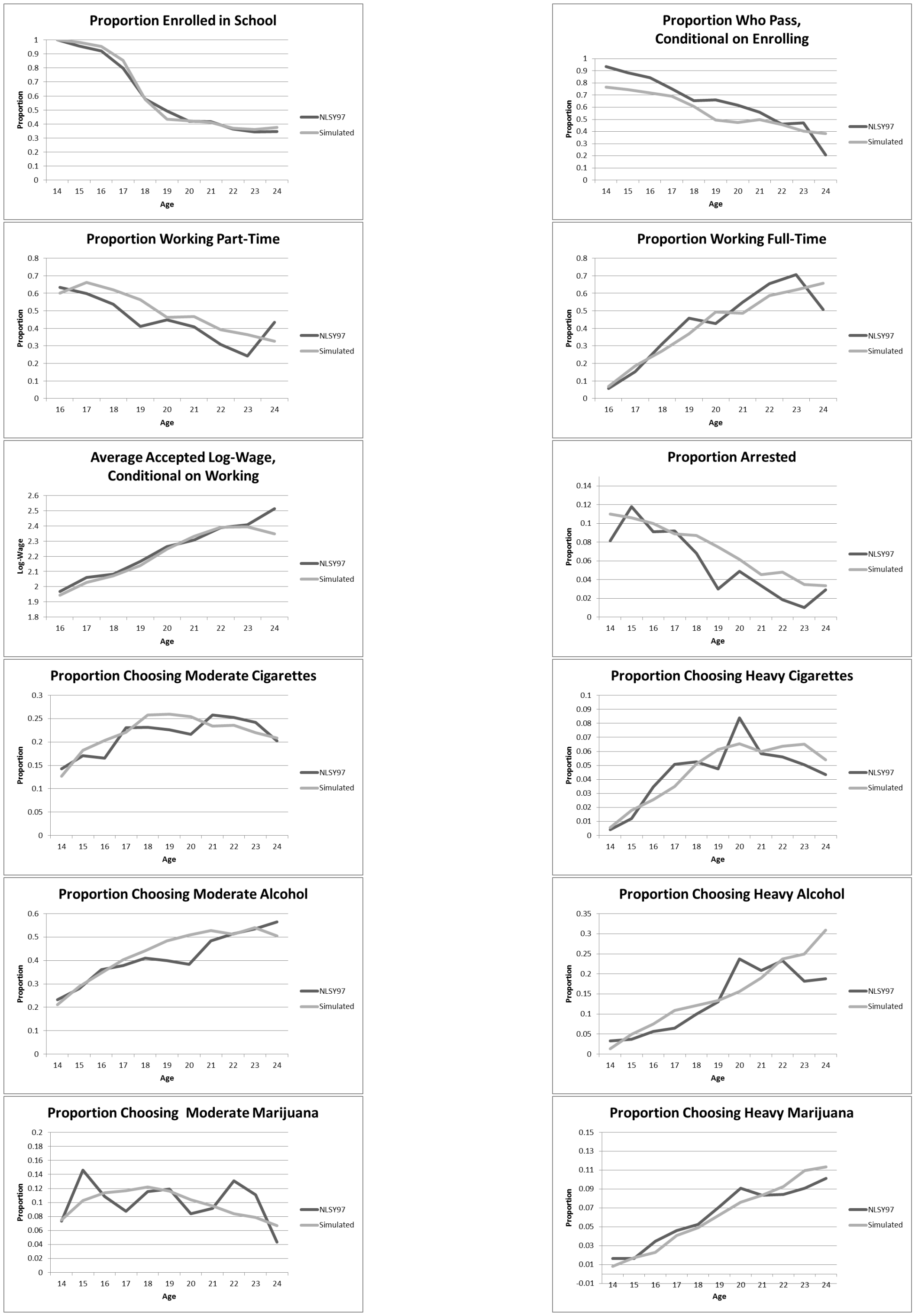
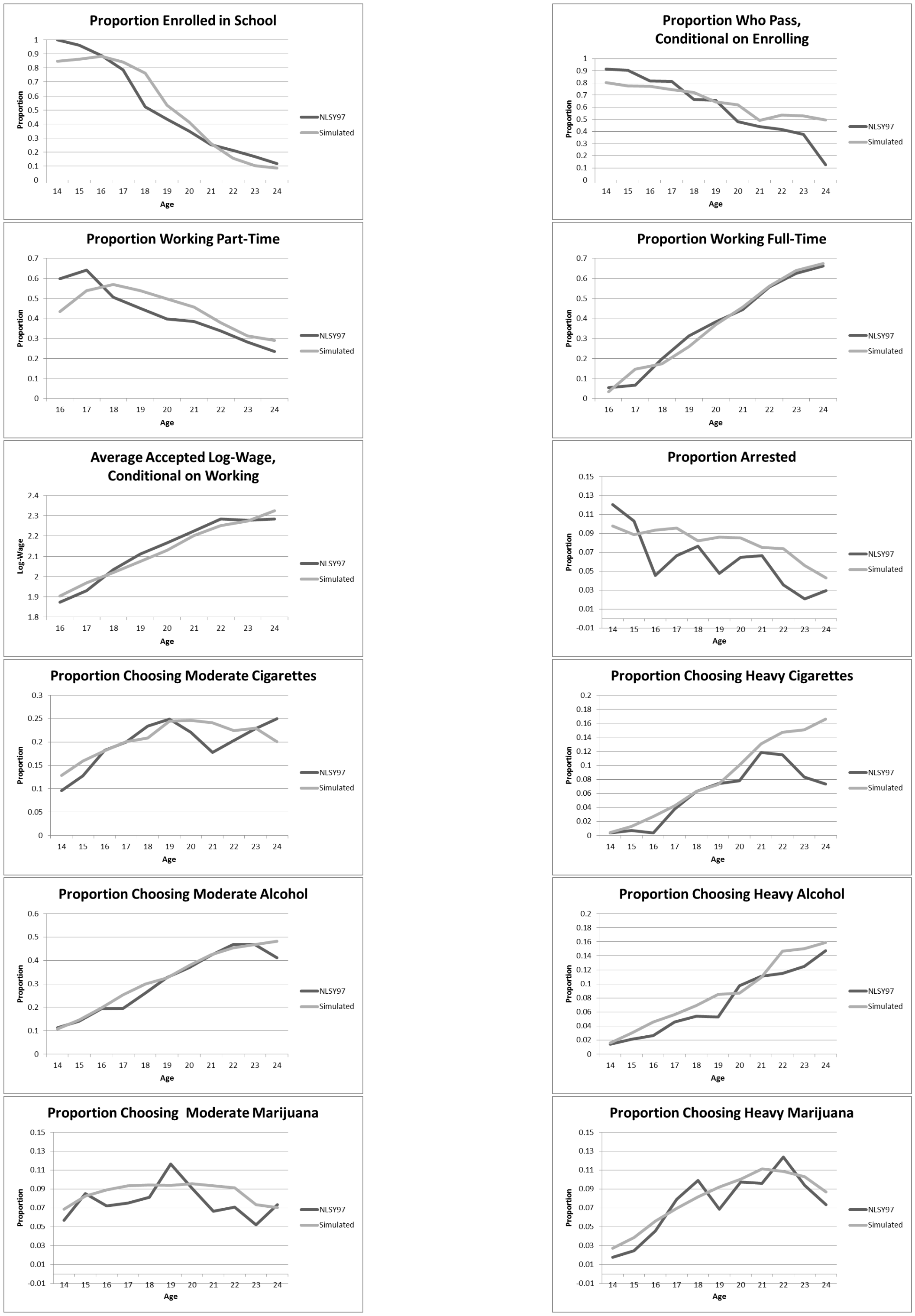


Figure 5: Comparison of Simulated and Actual Choices for Black Males



would result from the altered model. For example, changes in the arrest probabilities of minority males may change the equilibrium wage distribution. These equilibrium effects are not considered, because modeling equilibrium is beyond the scope of this paper.

8.1 The Role of Unobserved Preferences

The first counterfactual simulation demonstrates the importance of unobserved preferences on the choices to use substances. In this simulation, heterogeneity in preferences is eliminated by setting the time-persistent individual heterogeneity in productivity μ_i^w and in preferences μ_i^N equal to zero. In this simulation, all individuals have the same mean preferences for school enrollment, employment, and substance use. In this setting, substance use is driven entirely by nonpecuniary utility shocks, addiction, and gateway and complementarity effects.

Results of the baseline and counterfactual simulations are presented in Figures 6 through 8. Cigarette and marijuana use are lower for all males without heterogeneity in preferences, suggesting that individuals choose to use cigarettes and marijuana because of heterogeneity in preferences. Preferences do not play as important of a role in the alcohol use of Hispanic males, but eliminating heterogeneity in preferences does decrease the proportion of white and black males choosing to use a heavy amount of alcohol. Interestingly, although substance use decreases and the proportion of males being arrested decreases, there is very little change in enrollment and employment choices of males. The only exception is that Hispanic males are slightly more likely to enroll in school; enrollment increases by approximately 4 percentage points each year. This suggests that substance use is not causing poor labor market outcomes.

Figure 6: Comparison of Choices With and Without Heterogeneity in Preferences for White Males

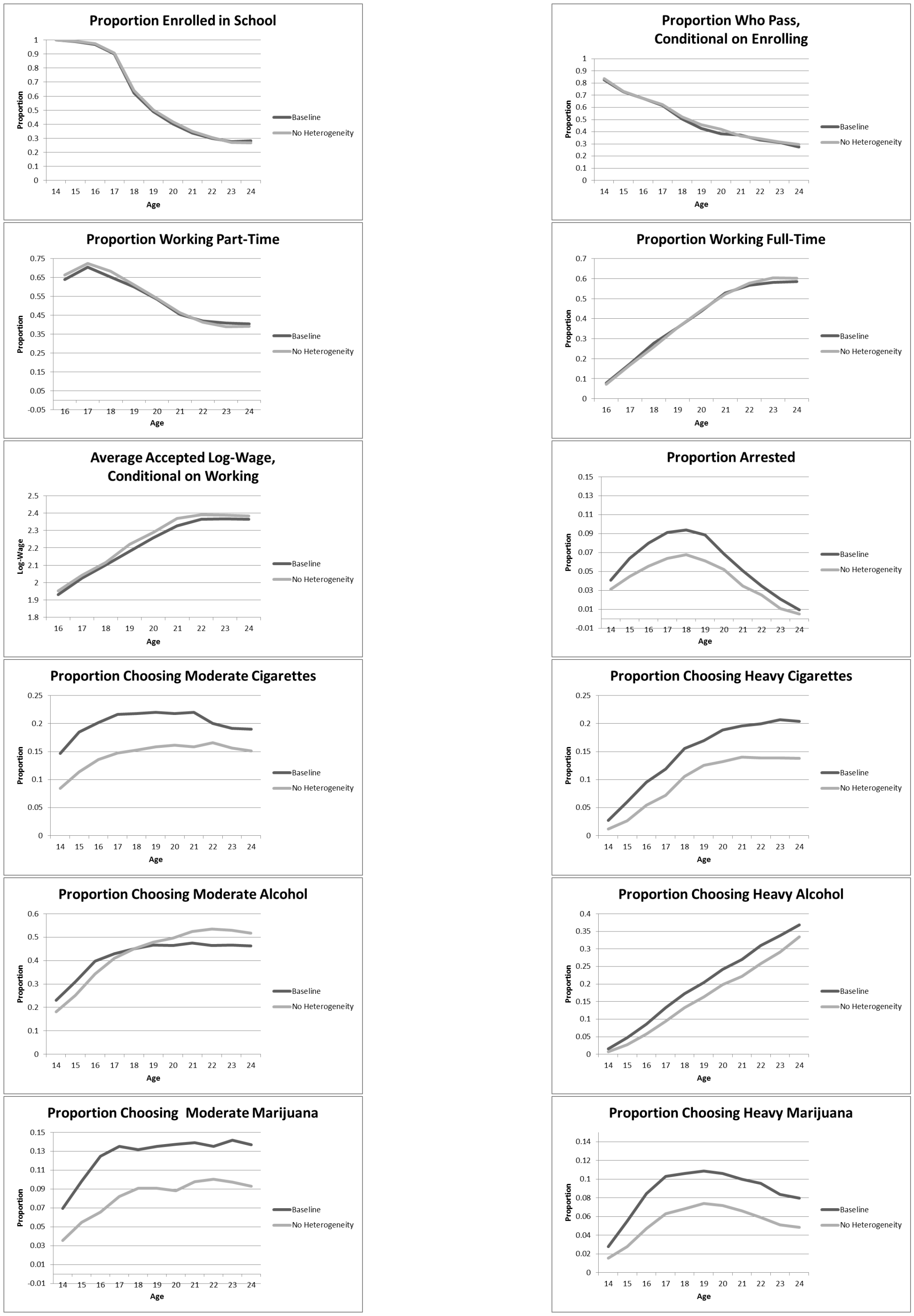


Figure 7: Comparison of Choices With and Without Heterogeneity in Preferences for Hispanic Males

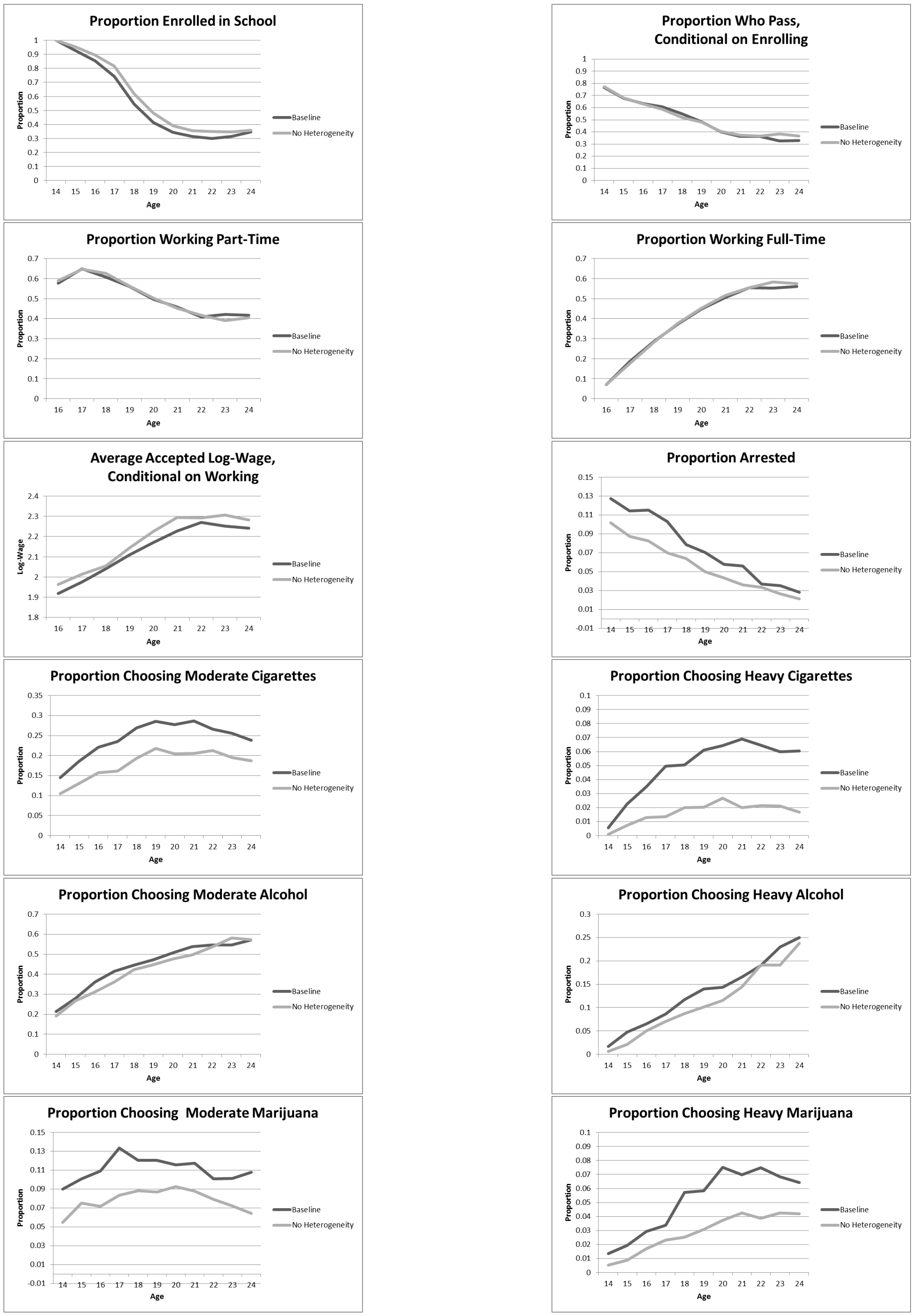
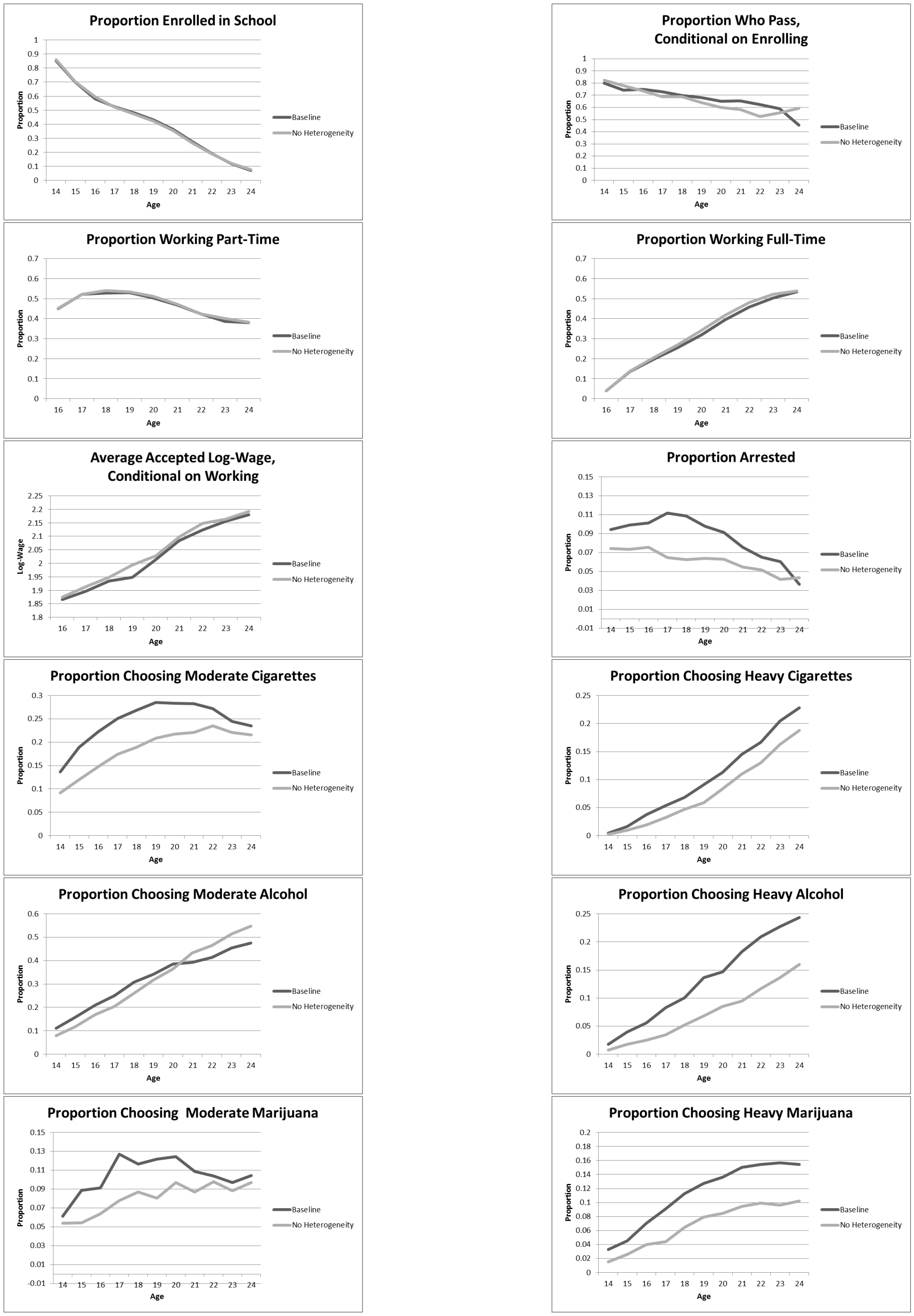


Figure 8: Comparison of Choices With and Without Heterogeneity in Preferences for Black Males



8.2 Gateway Effects

In this subsection, I present simulations that highlight the interdependencies of the three substances. In order to do this, for each substance, I evaluate the effects of turning off the gateway and complementarity effects of one or both of the other substances. For example, for cigarettes, if I turn off the interactions with past and present alcohol and marijuana use in the nonpecuniary utility of cigarettes, then cigarette use is driven solely by nonpecuniary utility shocks and addiction. On the other hand, if I only turn off the interactions with past and present marijuana use, cigarette use is additionally driven by gateway and complementarity effects of alcohol.

Results are presented in Figures 9 through 11. Comparing the baseline simulations with the simulations with no gateway effects shows that gateway effects play an important role in the choices to use cigarettes and marijuana. Without gateway and complementarity effects from the other substances, far fewer individuals use cigarettes or marijuana. Gateway effects are affecting individuals' decisions to use alcohol, but the effect is smaller. For white and black males, alcohol use has no effect on the use of heavy cigarettes, whereas marijuana use does. Only including the marijuana gateway effects explains almost the entire difference between the baseline simulation and the simulation with no gateway effects. Marijuana use of all males, on the other hand, is more influenced by alcohol gateway effects than cigarette gateway effects. However, neither effect fully explains the difference between the baseline and the simulation with no gateway effects. This suggests that the combination of cigarette and alcohol use is influencing marijuana use, rather than one substance alone.

Figure 9: Comparison of Choices With and Without Gateway Effects for White Males

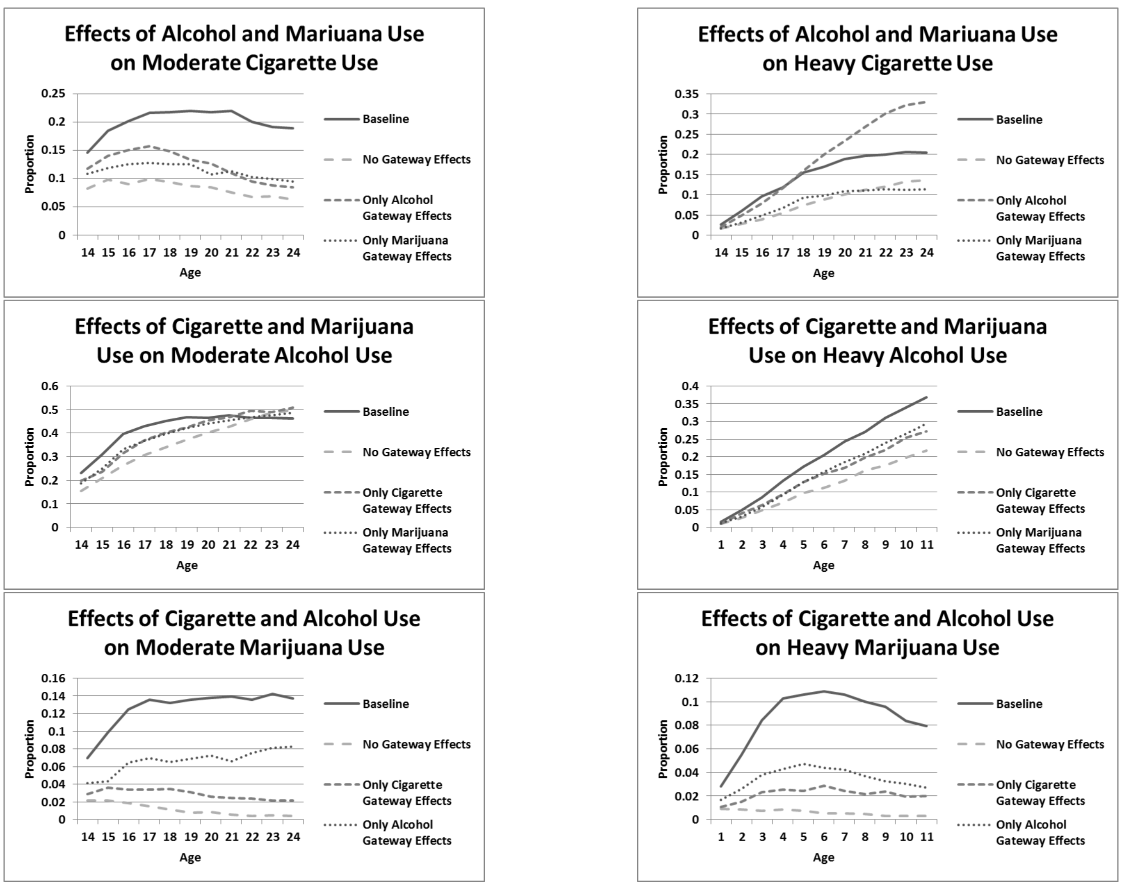


Figure 10: Comparison of Choices With and Without Gateway Effects for Hispanic Males

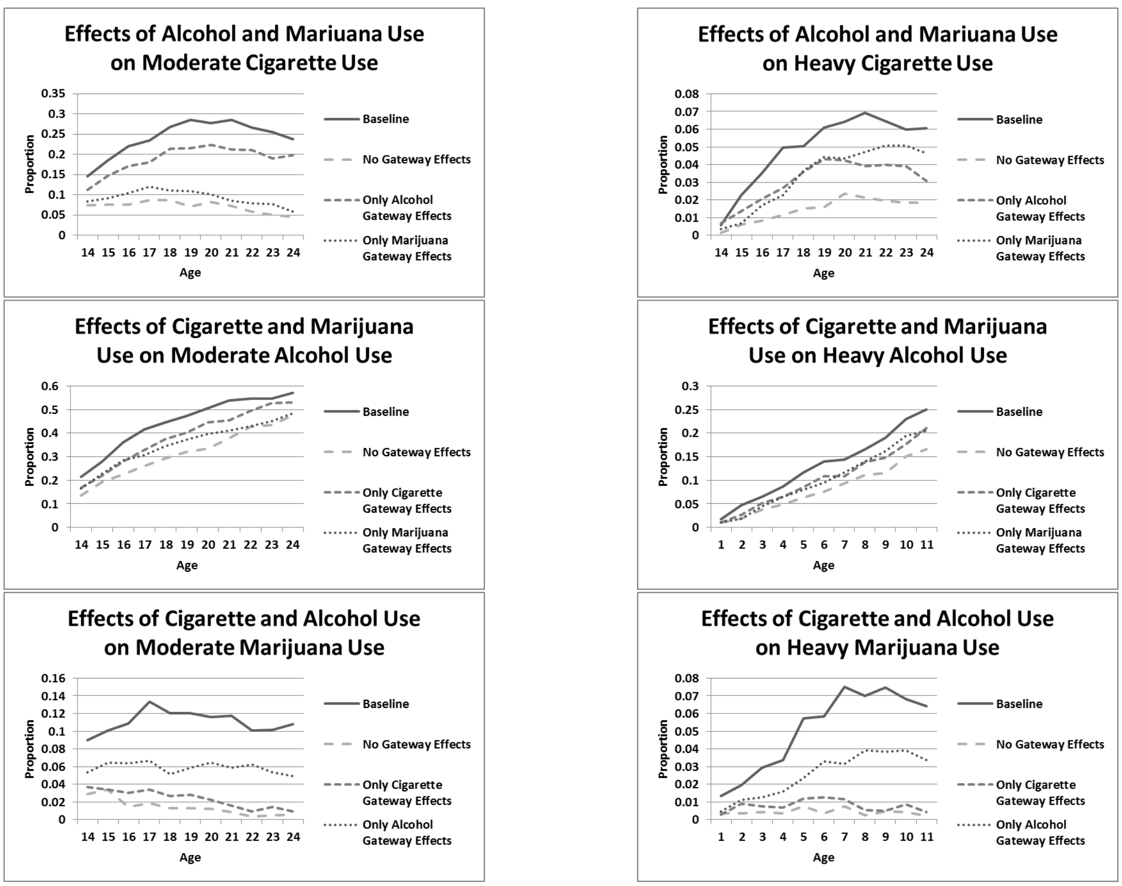
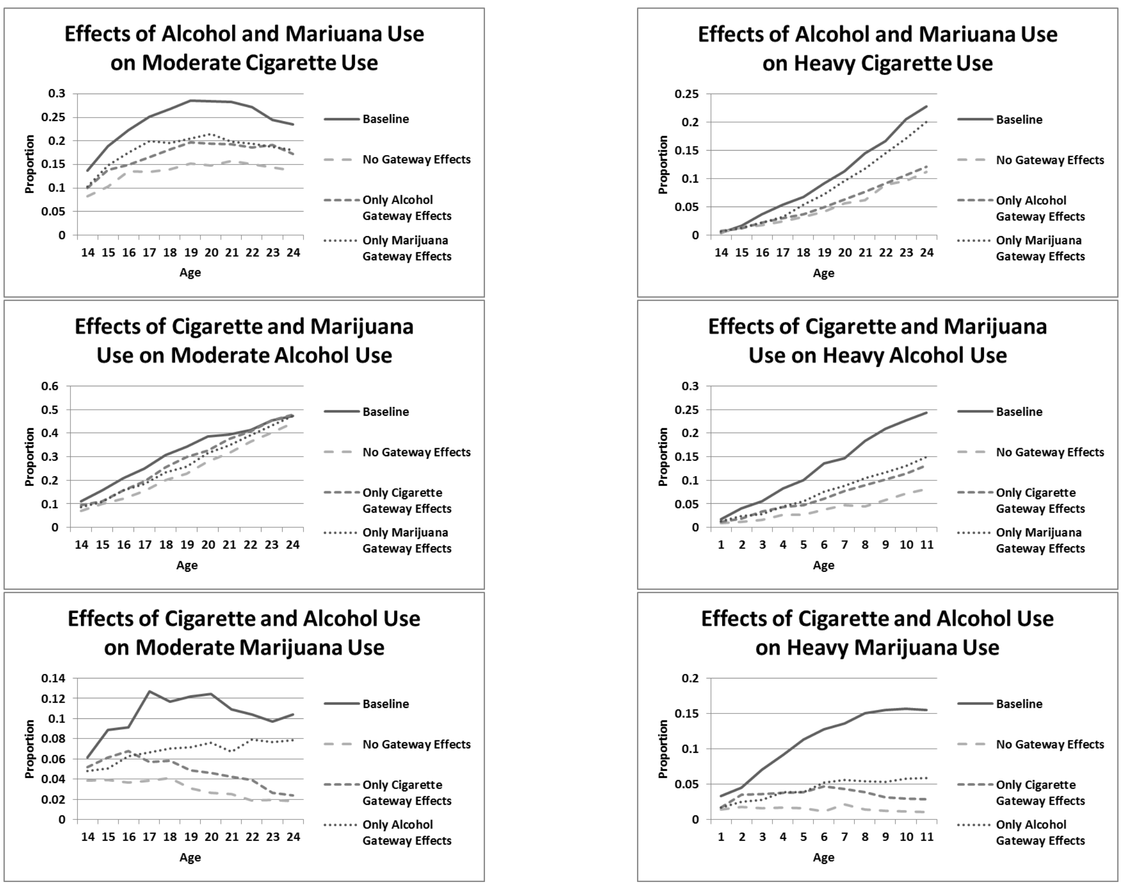


Figure 11: Comparison of Choices With and Without Gateway Effects for Black Males



8.3 Changing the Probability of Arrest of Minority Males

Lastly, I present the results of two policy simulations aimed at improving outcomes of minority males by changing their probability of arrest. White males are less likely to get arrested than black or Hispanic males. They are also more likely to graduate from high school and to be working full-time at the age of 24, even though their substance use is comparable to the use of black and Hispanic males. In the first policy experiment, I explore the outcome of setting the coefficients in the probability of arrest for black and Hispanic males equal to that of white males. Results are presented in Table 14.

Table 14 compares outcomes for minority males at the age of 24. The average number of arrests accumulated by the age of 24 decrease by .22 arrests for Hispanic males and .32 arrests for black males. High school graduation rates increase by 2.4 percentage points and 2.9 percentage points for Hispanic and black males, respectively; though the effect for Hispanic males is only significant at the 10 percent level. There is no effect of changing the probability of arrest on the cigarette or marijuana use of Hispanic males, whereas the proportion of Hispanic males using moderate alcohol decreases and the proportion using heavy alcohol increases. Two effects are causing this increase in heavy alcohol consumption. First, the marginal effect of using heavy alcohol on the probability of arrest decreases under the new policy. Additionally, heavy alcohol use and school enrollment are complements, as shown by the positive coefficient on heavy alcohol use presented in Table 5. Decreasing the probability of arrest increases enrollment, which causes more individuals to use alcohol. The effect of the policy change on the substance use of black males is quite different. I find there is no effect on the proportion of black males using moderate amounts of substances; however, the policy decreases the proportion of individuals using heavy cigarettes by 2.5 percentage points (11 percent), the proportion using heavy alcohol by 3.4 percentage points (14 percent), and the proportion using heavy marijuana by 4.3 percentage points (28 percent). This effect is coming from a decrease in the proportion of individuals using heavy amounts of marijuana at younger ages. While, the marginal effect of heavy marijuana on the probability of arrest decreases, which should encourage black males to consume more, using heavy marijuana decreases the probability of passing a grade conditional on enrollment. More individuals are enrolling, so the effect on passing is causing individuals to use less heavy marijuana at younger ages. Due to the gateway and complementarity effects of marijuana use on cigarette and alcohol use presented in the previous subsection, heavy cigarette and alcohol use also decrease.

Table 14: Effect of Counterfactual Experiment #1 on Outcomes of Minority Males at Age 24

Outome at Age 24	Hispanic Males			Black Males		
	Baseline	Counterfactual #1	T-Test	Baseline	Counterfactual #1	T-Test
	Mean	Mean	Difference (p-value)	Mean	Mean	Difference (p-value)
High School Graduation Rate	0.633	0.657	-0.024 (0.08)	0.432	0.461	-0.029* (0.03)
Proportion Working Part-Time	0.416	0.401	0.015 (0.27)	0.380	0.393	-0.013 (0.31)
Proportion Working Full-Time	0.560	0.584	-0.024 (0.09)	0.534	0.514	0.020 (0.13)
Accepted Log-Wage	2.241	2.275	-0.034* (0.01)	2.180	2.181	-0.001 (0.95)
Number of Arrests	0.822	0.599	0.224* (0.00)	0.942	0.626	0.316* (0.00)
Proportion Using Moderate Cigarettes	0.238	0.234	0.004 (0.74)	0.235	0.243	-0.008 (0.49)
Proportion Using Heavy Cigarettes	0.061	0.052	0.009 (0.17)	0.228	0.203	0.025* (0.02)
Proportion Using Moderate Alcohol	0.570	0.531	0.039* (0.01)	0.474	0.476	-0.002 (0.89)
Proportion Using Heavy Alcohol	0.250	0.278	-0.028* (0.03)	0.243	0.210	0.034* (0.00)
Proportion Using Moderate Marijuana	0.108	0.111	-0.003 (0.75)	0.104	0.096	0.008 (0.33)
Proportion using Heavy Marijuana	0.064	0.052	0.012 (0.07)	0.155	0.112	0.043* (0.00)

* p<0.05.

In my second policy experiment, I simulate the effects of legalizing marijuana on black males by setting the coefficients on moderate and heavy marijuana in the probability of arrest equation presented in Table 10 equal to the coefficients on moderate and heavy alcohol. I believe this is a reasonable approximation of the lower bound of the effect of legalizing marijuana since both substances have the same age restriction and are used for similar purposes. I believe this is the lower bound because individuals tend to commit more crimes when intoxicated by alcohol rather than when intoxicated by marijuana; therefore, it is likely that legalizing marijuana would decrease the effect of marijuana use on the probability of arrest by more that what I am decreasing it by in this simulation.[20] This experiment is not appropriate to run on white or Hispanic males because the coefficients on alcohol are higher than those on marijuana and I do not believe that legalizing marijuana will increase the probability of arrest. Results are presented in Table 15. I find that the average number of arrests accumulated by the age of 24 decreases by .164 arrests. There is no effect on high school graduation rates. I again find that decreasing the probability of arrest decreases the proportion of individuals using heavy cigarettes by 10.1 percent, heavy alcohol by 10.7 percent, and heavy marijuana by 31 percent.

9 Conclusion

In this paper, I develop a dynamic structural model in which individuals make decisions about schooling and work as well as how much alcohol, cigarettes, and marijuana to consume. My model allows substance use to affect career paths through its effects on

[20] According to the Federal Bureau of Investigation's Uniform Crime Reporting Program, in 2012 12.6 percent of crimes against persons were committed with alcohol involvement whereas only 2.1 percent were committed with drug involvement. Similarly, 1.5 percent of crimes against property were committed while using alcohol whereas only 0.9 percent were committed while using other drugs. Crimes against persons include crimes such as assault, homicide, kidnapping, and forcible and nonforcible sex offenses. Crimes against property include arson, burglary, forgery, vandalism, larceny, and robbery. More information can be found at: http://www.fbi.gov/about-us/cjis/ucr/nibrs/2012/table-pdfs/drugs-narcotics-and-alcohol-involvement-by-offense-category-2012

Table 15: Effect of Counterfactual Experiment #2 on Outcomes of Black Males at Age 24

Outome at Age 24	Baseline	Counterfactual #1	T-Test
	Mean	Mean	Difference (p-value)
High School Graduation Rate	0.432	0.439	-0.007 (0.59)
Proportion Working Part-Time	0.380	0.367	0.013 (0.32)
Proportion Working Full-Time	0.534	0.529	0.005 (0.69)
Accepted Log-Wage	2.180	2.173	0.007 (0.55)
Number of Arrests	0.942	0.778	0.164* (0.00)
Proportion Using Moderate Cigarettes	0.235	0.235	0.000 (0.97)
Proportion Using Heavy Cigarettes	0.228	0.205	0.023* (0.04)
Proportion Using Moderate Alcohol	0.474	0.474	0.000 (0.98)
Proportion Using Heavy Alcohol	0.243	0.217	0.026* (0.02)
Proportion Using Moderate Marijuana	0.104	0.118	-0.014 (0.09)
Proportion using Heavy Marijuana	0.155	0.107	0.048* (0.00)

* $p<0.05$.

educational attainment and wage offers. Individuals are rational and forward-looking and make decisions in order to maximize their discounted lifetime utility, accounting for the effects of substance use on future outcomes. I improve on the current literature by allowing individuals in my model to make decisions about their level of alcohol, cigarettes, and marijuana use. This allows me to differentiate between the effects of moderate versus heavy substance use. It also allows me to capture the interdependencies of cigarette, alcohol, and marijuana use.

I estimate my model using Bayesian Markov Chain Monte Carlo methods. Using classical estimation techniques to estimate my model is difficult for several reasons. It is important in my model that I allow for a substantial amount of time-persistent unobserved heterogeneity in order to identify the causal effects of state dependence on outcomes. MCMC estimation offers the convenience of data augmentation that allows me to incorporate this unobserved heterogeneity into my model. Additionally, estimation with classical methods often requires evaluating the value function for each sample observation and for each trial guess of the parameters. This can be very time consuming in problems with large choice sets and large state spaces. The IJC algorithm provides a way to approximate the value functions by using stored value functions from earlier iterations of the MCMC algorithm, making estimation of my model feasible.

I find that heterogeneity in preferences plays an important role in an individual's choice to use cigarettes and marijuana. That is, without heterogeneity, the proportion of individuals using cigarettes and marijuana would be less. I find that cigarette and alcohol use decrease the wages of white and Hispanic males, but have no statistically significant effect on the wages of black males. Marijuana use has no significant effect on wages of any males. I find that past and present alcohol and cigarette use increase the probability that an individual uses marijuana, suggesting that alcohol, cigarettes, and marijuana are complements. Additionally, my results suggest that a policy maker trying to decrease marijuana use may benefit from not only trying to decrease the use of either cigarettes or alcohol of young males, but by trying to decrease both simultaneously. I also find evidence supporting the reverse gateway theory. That is, I find that marijuana use leads individuals into being heavy cigarette users. Based on my results, it is unclear how legalizing marijuana will affect cigarette use. If past marijuana use increases cigarette use because individuals are substituting away from marijuana at a certain age, then legalizing marijuana will not increase cigarette use. If, on the other hand, past marijuana use increases cigarette use because individuals like to use the two substances together, then legalizing marijuana may increase cigarette use. More work needs to be done to distinguish which reason is causing marijuana use to be a gateway into cigarette use. This finding will be important for policy makers to consider as states continue legalizing the recreational use of marijuana.

Lastly, I find that white males have a lower probability of arrest than black and Hispanic males, conditional on age, previous arrests, and substance use. White males are also more likely to graduate from high school and to be working a full-time job at the age of 24. This suggests that arrests may be contributing to the education gap between white and minority males. I run two policy simulations to see if the outcomes of minority males can be improved by decreasing the probability of arrest. In the first, I explore what would happen if Hispanic and black males faced the same probability of arrest as white males, conditional on states and choices. I find that high school graduation rates increase by 3.8 percent and 6.7 percent for Hispanic and black males, respectively. Additionally, the proportion of black males using heavy cigarettes at the age of 24 decreases by 11 percent, the proportion using heavy alcohol decreases by 14 percent, and the proportion using heavy marijuana decreases by 28 percent.

In my main results, I also find that black males are disproportionately more likely to get arrested because of marijuana use. In my second policy experiment, I consider the effects of decreasing the marginal effects of marijuana use on the probability of arrest of black males, through, for example, legalizing marijuana. In this policy simulation, I set the coefficients on moderate and heavy marijuana use in the probability of arrest equation equal to the coefficients on moderate and heavy alcohol. This policy change has no effect on high school graduation, but I find it decreases the proportion of black males using heavy amounts of cigarettes, alcohol, and marijuana. These policy simulations suggest that policing practices that result in increased arrests of minority males may have counterproductive results, as they contribute to increased substance use and reduced education.

References

[1] James H. Albert and Siddhartha Chib. Bayesian Analysis of Binary and Polychotomous Response Data. *Journal of American Statistical Association*, 88(422):669–679, 1993.

[2] Catherine Alford. Medical Marijuana Laws and the Market for Marijuana. 2013.

[3] D Mark Anderson, Benjamin Hansen, and Daniel I Rees. Medical marijuana laws, traffic fatalities, and alcohol consumption. *Journal of Law and Economics*, 56(2):333–369, 2013.

[4] Peter Arcidiacono, Holger Sieg, and Frank Sloan. LIVING RATIONALLY UNDER THE VOLCANO? AN EMPIRICAL ANALYSIS OF HEAVY DRINKING AND SMOKING*. *International Economic Review*, 48(1):37–65, 2007.

[5] Gary S. Becker, Michael Grossman, and Kevin M. Murphy. An Empirical Analysis of Cigarette Addiction. *The American Economic Review*, 84(3):396–418, 1994.

[6] Gary S Becker and Kevin M Murphy. A Theory of Rational Addiction. *Journal of Political Economy*, 96(4):pp. 675–700, 1988.

[7] Jeremy W Bray. Alcohol Use, Human Capital, and Wages. *Journal of Labor Economics*, 23(2):pp. 279–312, 2005.

[8] Jeremy W Bray, Gary A Zarkin, Chris Ringwalt, and Junfeng Qi. The relationship between marijuana initiation and dropping out of high school. *Health Economics*, 9(1):9–18, 2000.

[9] Thomas C Buchmueller and Samuel H Zuvekas. Drug use, drug abuse, and labour market outcomes. *Health Economics*, 7(3):229–245, 1998.

[10] George Casella and Edward I George. Explaining the gibbs sampler. *The American Statistician*, 46(3):167–174, 1992.

[11] E.F. Cataldi, C. Green, R. Henke, T. Lew, J. Woo, B. Sepherd, and P. Siegal. 2008-09 baccalaureate and beyond longitudinal study: First look (nces 2011-236). *U.S. Department of Education. Washington, DC: National Center for Education Statistics.*, 2011.

[12] Frank Chaloupka. Rational Addictive Behavior and Cigarette Smoking. *The Journal of Political Economy*, 99(4):722–742, 1991.

[13] Frank J Chaloupka and Adit Laixuthai. Do youths substitute alcohol and marijuana? some econometric evidence. *Eastern Economic Journal*, pages 253–276, 1997.

[14] Frank J Chaloupka, Rosalie Liccardo Pacula, Matthew C Farrelly, Lloyd D Johnston, and Patrick M O'Malley. Do higher cigarette prices encourage youth to use marijuana? Technical report, National bureau of economic research, 1999.

[15] Pinka Chatterji. Illicit drug use and educational attainment. *Health Economics*, 15(5):489–511, 2006.

[16] AndrewT. Ching, Susumu Imai, Masakazu Ishihara, and Neelam Jain. A practitioner's guide to Bayesian estimation of discrete choice dynamic programming models. *Quantitative Marketing and Economics*, 10(2):151–196, 2012.

[17] Phillip J Cook and Michael J Moore. Drinking and schooling. *Journal of Health Economics*, 12(4):411–429, 1993.

[18] Washington Student Achievement Council. National tuition and fee report. Various Volumes, 1996-2010.

[19] Benjamin Crost and Daniel I. Rees. The minimum legal drinking age and marijuana use: New estimates from the {NLSY97}. *Journal of Health Economics*, 32(2):474 – 476, 2013.

[20] Thomas S Dee. The complementarity of teen smoking and drinking. *Journal of Health Economics*, 18(6):769 – 793, 1999.

[21] Thomas S Dee and William N Evans. Teen Drinking and Educational Attainment: Evidence from Two-Sample Instrumental Variables Estimates. *Journal of Labor Economics*, 21(1):pp. 178–209, 2003.

[22] Jeff DeSimone. Illegal Drug Use and Employment. *Journal of Labor Economics*, 20(4):952–977, 2002.

[23] John DiNardo and Thomas Lemieux. Alcohol, marijuana, and american youth: the unintended consequences of government regulation. *Journal of Health Economics*, 20(6):991 – 1010, 2001.

[24] Zvi Eckstein and Kenneth I Wolpin. Why Youths Drop out of High School: The Impact of Preferences, Opportunities, and Abilities. *Econometrica*, 67(6):1295–1339, 1999.

[25] Matthew C. Farrelly, Jeremy W. Bray, Gary A. Zarkin, and Brett W. Wendling. The joint demand for cigarettes and marijuana: evidence from the national household surveys on drug abuse. *Journal of Health Economics*, 20(1):51 – 68, 2001.

[26] Alan E Gelfand and Adrian FM Smith. Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, 85(410):398–409, 1990.

[27] Stuart Geman and Donald Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6):721–741, 1984.

[28] John Geweke. Efficient simulation from the multivariate normal and student-t distributions subject to linear constraints and the evaluation of constraint probabilities. In *Computing Science and Statistics: Proceedings of the ...*, pages 1–14, 1991.

[29] John F Geweke, Michael P Keane, and David E Runkle. Statistical inference in the multinomial multiperiod probit model. *Journal of Econometrics*, 80(1):125–165, 1997.

[30] Donna B. Gilleskie and Koleman S. Strumpf. The Behavioral Dynamics of Youth Smoking. *The Journal of Human Resources*, 40(4):822–866, 2005.

[31] Jonathan Gruber and Botond Köszegi. Is Addiction “Rational”? Theory and Evidence. *The Quarterly Journal of Economics*, 116(4):1261–1303, 2001.

[32] Lingxin Hao, V Joseph Hotz, and Ginger Z Jin. Games parents and adolescents play: Risky behaviour, parental reputation and strategic transfers*. *The Economic Journal*, 118(528):515–555, 2008.

[33] Lana D. Harrison, Steven S. Martin, Tihomir Enev, and Deborah Harrington. Comparing drug testing and self-report of drug use among youths and young adults in the

general population. Technical report, Substance Abuse and Mental Health Services Administration, Office of Applied Studies, Rockville, MD, 2007.

[34] James J Heckman and Bo E Honore. The empirical content of the roy model. *Econometrica: Journal of the Econometric Society*, pages 1121–1149, 1990.

[35] Kosuke Imai and David A van Dyk. A Bayesian analysis of the multinomial probit model using marginal data augmentation. *Journal of Econometrics*, 124(2):311–334, 2005.

[36] Susumu Imai, Neelam Jain, and Andrew Ching. Bayesian Estimation of Dynamic Discrete Choice Models. *Econometrica*, 77(6):1865–1899, 2009.

[37] Mireille Jacobson. Baby booms and drug busts: Trends in youth drug use in the United States, 1975-2000. *Quarterly Journal of Economics*, 119:1481–1512, 2004.

[38] Robert Kaestner. The Effect of Illicit Drug Use on the Labor Supply of Young Adults. *The Journal of Human Resources*, 29(1):pp. 126–155, 1994.

[39] Michael P Keane and Kenneth I Wolpin. The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence. *The Review of Economics and Statistics*, 76(4):648–672, 1994.

[40] Michael P Keane and Kenneth I Wolpin. The Career Decisions of Young Men. *Journal of Political Economy*, 105(3):pp. 473–522, 1997.

[41] Melissa J. Krauss, Patricia A. Cavazos-Rehg, Andrew D. Plunk, Laura J. Bierut, and Richard A. Grucza. Effects of state cigarette excise taxes and smoke-free air policies on state per capita alcohol consumption in the united states, 1980 to 2009. *Alcoholism: Clinical and Experimental Research*, 38(10):2630–2638, 2014.

[42] José M Labeaga. A double-hurdle rational addiction model with heterogeneity: Estimating the demand for tobacco. *Journal of Econometrics*, 93(1):49–72, 1999.

[43] Ignacio Martinez. A Principal-Agent Model of Parental Transfers and Human Capital Accumulation. 2014.

[44] Robert McCulloch and Peter E Rossi. An exact likelihood analysis of the multinomial probit model. *Journal of Econometrics*, 64(1-2):207–240, September 1994.

[45] Alvaro Mezza. Illegal Drugs, Education, and Labor Market Outcomes. 2011.

[46] William Orzechowski and Robert Walker. The tax burden on tobacco. *Historical compilation*, 46, 2011.

[47] Matthew Osborne. Consumer learning, switching costs, and heterogeneity: A structural examination. *Quantitative Marketing and Economics*, 9(1):25–70, 2011.

[48] Rosalie Liccardo Pacula. Does increasing the beer tax reduce marijuana consumption? *Journal of Health Economics*, 17(5):557 – 585, 1998.

[49] Gabriel A. Picone, Frank Sloan, and Justin G. Trogdon. The effect of the tobacco settlement and smoking bans on alcohol consumption. *Health Economics*, 13(10):1063–1080, 2004.

[50] Mark R Rosenzweig and Kenneth I Wolpin. Parental and public transfers to young women and their children. *The American Economic Review*, pages 1195–1212, 1994.

[51] Andrew Roy. Some Thoughts on the Distribution of Earnings. *Oxfor Economic Papers*, 3:135–146, 1951.

[52] Christopher J. Ruhm, Allison Snow Jones, Kerry Anne McGeary, William C. Kerr, Joseph V. Terza, Thomas K. Greenfield, and Ravi S. Pandian. What U.S. data should be used to measure the price elasticity of demand for alcohol? *Journal of Health Economics*, 31(6):851–862, 2012.

[53] John Rust. Using Randomization to Break the Curse of Dimensionality. *Econometrica*, 65(3):487–516, 1997.

[54] Adrian FM Smith and Gareth O Roberts. Bayesian computation via the gibbs sampler and related markov chain monte carlo methods. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 3–23, 1993.

[55] Paul Sullivan. A DYNAMIC ANALYSIS OF EDUCATIONAL ATTAINMENT, OCCUPATIONAL CHOICES, AND JOB SEARCH*. *International Economic Review*, 51(1):289–317, 2010.

[56] Clifford F Thies and Charles A Register. Decriminalization of marijuana and the demand for alcohol, marijuana and cocaine. *The Social Science Journal*, 30(4):385–399, 1993.

[57] Luke Tierney. Markov chains for exploring posterior distributions. *the Annals of Statistics*, pages 1701–1728, 1994.

[58] High Times. Trans-high market quotations. *New York: Trans-High Corporation Press*, Various Issues, 1996-2010.

[59] Kenneth Train. *Discrete Choice Methods with Simulation*. Cambridge University Press, second edition, 2009.

[60] J. Williams, Rosalie Liccardo Pacula, Frank J. Chaloupka, and Henry Wechsler. Alcohol and marijuana use among college students: economic complements or substitutes? *Health Economics*, 13(9):825–843, 2004.

[61] Tetsuji Yamada, Michael Kendix, and Tadashi Yamada. The impact of alcohol consumption and marijuana use on high school graduation. *Health Economics*, 5(1):77–92, 1996.

[62] Gary A Zarkin, Michael T. French, Thomas Mroz, and Jeremy W Bray. Alcohol use and wages: New results from the national household survey on drug abuse. *Journal of Health Economics*, 17(1):53–68, 1998.

[63] Yiyi Zhou. Bayesian Estimation of a Dynamic Model of Oligopoly Pricing and Entry with Strategic Consumers: Application to Video Games. 2011.

A Data

A.1 Auxiliary Tables

Table 16: Comparison of Substance Use Rates of 12-17 Year-Olds in 1997

	Cigarette		**Beer**		**Marijuana**	
	NLSY97	NHSDA	NLSY97	NHSDA	NLSY97	NHSDA
Total Sample	0.417	0.387	0.450	0.397	0.210	0.189
Female	0.417	0.383	0.442	0.407	0.198	0.182
Male	0.417	0.390	0.458	0.388	0.221	0.195
White	0.445	0.421	0.476	0.425	0.215	0.196
Black, non-Hispanic	0.315	0.282	0.354	0.314	0.183	0.161
Hispanic	0.384	0.330	0.429	0.363	0.208	0.167

This table compares self-reported substance use rates in the 1997 National Longitudinal Survey of Youth
with those found in the 1997 National Household Survey on Drug Abuse.

Table 17: Alcohol Price Imputation Regression

	b/se
State Beer Tax	1.2315***
	(0.21)
Post 2000	2.7718***
	(0.05)
Constant	3.2085***
	(0.23)
N	583
R^2	0.971

The dependent variable is equal to the average price in state s and year t and grade g. I control for state and year fixed effects. I include Post 2000 as a regressor to adjust for the change in the type of beer sampled after 2000. Standard errors are in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Table 18: Marijuana Price Imputation Regression

	b/se
Medical Marijuana Law	-29.4921**
	(13.25)
Medical Marijuana Law and Decriminalized	-2.3898
	(21.81)
Marijuana Decriminalized	11.4161
	(36.82)
Murder Rate	-0.5604
	(2.72)
Property Crime Rate	0.0319***
	(0.01)
Violent Crime Rate	-0.0883*
	(0.05)
Low-Grade Marijuana	-247.0313***
	(4.71)
Constant	266.9717***
	(48.14)
N	887
R^2	0.805

The dependent variable is equal to the average price in state s and year t and grade g. I control for state and year fixed effects. Crime Rates are the rate per 100,000 residents. Standard errors are in parentheses.
* $p<0.10$, ** $p<0.05$, *** $p<0.01$

A.2 Descriptive Statistics

I run several regressions to see how past substance use is associated with current substance use, work, and education outcomes at age 23. None of the relationships I describe below should be interpreted as causal, but they do describe key patterns in the data. The main concern is that the enjoyment that an individual gets from substance use is likely correlated with the educational and employment outcomes of interest for other reasons. The measures of past substance use that I use in these regressions are the total number of years the individual chose to consume any amount of a substance.

I first look to see how past substance use affects substance use at age 23. Table 19 presents the results of an ordered probit where the dependent variable is the choice of substance use at age 23; individuals can choose no use, moderate use, or heavy use.[21] As in Table 3, marginal effects predict the probability of an individual *not* using. For all substances, past use increases the probability of using that substance at age 23. However, past cigarette, alcohol, or marijuana use does not affect the probability of using the other substances.

Next, I see how past substance use affects the highest grade attained by age 23. To do this, I run a regression of highest grade attained at age 23 on past substance use and the cost of college. Results are presented in Table 20. Here, there are two regressions: one that excludes arrests and one that includes arrests. In the first regression, past cigarette use decreases educational attainment, while alcohol use increases it. Next, I add in arrests as a covariate and find that past marijuana use increases the highest grade attained by age 23, after controlling for arrests. This counterintuitive result may arise because those who use

[21] I do not present all of the marginal effects in this section. Please contact the author if you would like to see additional marginal effects.

Table 19: Ordered Probit: Substance Use at Age 23

	Cigarette		Alcohol		Marijuana	
	Estimates	Marginal Effects	Estimates	Marginal Effects	Estimates	Marginal Effects
Cigarette Stock	0.404***	-0.078***	-0.027	0.007	0.024	-0.004
	(0.03)	(0.00)	(0.02)	(0.01)	(0.03)	(0.01)
Alcohol Use Stock	0.002	0.000	0.262***	-0.068***	0.046	-0.008
	(0.03)	(0.00)	(0.03)	(0.01)	(0.04)	(0.01)
Marijuana Use Stock	0.007	-0.001	0.042	-0.011	0.325***	-0.055***
	(0.03)	(0.01)	(0.03)	(0.01)	(0.04)	(0.01)
Cigarette Price	-0.003	0.001	-0.081*	0.021*	-0.039	0.007
	(0.06)	(0.01)	(0.05)	(0.01)	(0.07)	(0.01)
Beer Price	-0.246*	0.048*	-0.117	0.030	-0.275	0.046
	(0.15)	(0.03)	(0.11)	(0.03)	(0.17)	(0.03)
Marijuana Price	0.003**	-0.001**	0.000	0.000	0.001	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Black	0.134	-0.026	-0.107	0.028	0.530**	-0.089**
	(0.18)	(0.04)	(0.14)	(0.04)	(0.22)	(0.04)
Hispanic	-0.285	0.055	-0.160	0.042	0.442**	-0.074**
	(0.18)	(0.03)	(0.14)	(0.04)	(0.19)	(0.03)
N	519		519		519	

Marginal effects predict the probability of an individual not using. Data come from the NLSY97. The sample is limited to those who were interviewed from ages 14 to 23. Cigarette price reflects the price of a pack of cigarettes. A pack generally contains 20 cigarettes. Beer price reflects the price of a 6-pack of beer. Marijuana price reflect the price of an ounce of marijuana. A joint of marijuana contains around half of a gram.
* p<0.10, ** p<0.05, *** p<0.01

Table 20: Highest Grade Completed by Age 23

	Estimates	Estimates
Cigarette Stock	-0.404***	-0.348***
	(0.04)	(0.04)
Alcohol Use Stock	0.252***	0.251***
	(0.04)	(0.04)
Marijuana Use Stock	0.061	0.105**
	(0.05)	(0.05)
Tuition Price	0.024	0.033
	(0.05)	(0.05)
Black	-0.922***	-0.827***
	(0.25)	(0.24)
Hispanic	-0.824***	-0.777***
	(0.23)	(0.23)
Total Arrests		-0.562***
		(0.09)
N	517	517

Data comes from the NLSY97. The sample is limited to those who were interviewed from ages 14 to 23. Tuition is in 1,000 of dollars.
* p<0.10, ** p<0.05, *** p<0.01

marijuana and do not get arrested also have some sort of unobserved characteristic that is correlated with higher education.

Next, Table 21 reports ordered probit estimates of how substance use affects full-time, part-time, and no work decisions. The marginal effects I present predict the probability of *not* working. Past cigarette use decreases the probability that an individual works. After controlling for arrests, marijuana use increases the probability that an individual works by 0.8 percentage points. Past alcohol use does not statistically significantly affect whether an individual works. Additionally, total arrests are not associated with the work decision.

Lastly, I present how past substance use affects the log-wage for individuals who are observed working. Results are presented in Table 22. Cigarette and marijuana use are not statistically significantly associated with wages. However, having used alcohol for one more year increases wages by 4 percent. This regression does not control for selection into working. As suggested in the previous regression, substance use may affect whether or

Table 21: Ordered Probit: Work Hours for Individuals at Age 23

	Estimates	Marginal Effects	Estimates	Marginal Effects
Work Experience	0.125***	-0.021***	0.125***	-0.021***
	(0.02)	(0.00)	(0.02)	(0.00)
Education	0.538**	0.007**	0.482**	0.010**
	(0.21)	(0.00)	(0.22)	(0.00)
Education Squared	-0.021***		-0.020**	
	(0.01)		(0.01)	
Cigarette Stock	-0.066***	0.011***	-0.061**	0.010**
	(0.02)	(0.00)	(0.02)	(0.00)
Alcohol Use Stock	-0.010	0.002	-0.006	0.001
	(0.02)	(0.00)	(0.02)	(0.00)
Marijuana Use Stock	0.038	-0.007	0.050*	-0.008*
	(0.03)	(0.00)	(0.03)	(0.00)
Unemployment Rate	-0.118***	0.020***	-0.117**	0.020**
	(0.05)	(0.01)	(0.05)	(0.01)
Black	-0.124	0.021	-0.125	0.021
	(0.14)	(0.02)	(0.14)	(0.02)
Hispanic	0.180	-0.031	0.177	-0.030
	(0.13)	(0.02)	(0.13)	(0.02)
Total Arrests			-0.087	0.016
			(0.11)	(0.02)
Total Arrests Squared			-0.010	
			(0.02)	
N	517		517	

Marginal effects predict the probability of an individual not working. Data come from the NLSY97. Data comes from the NLSY97. The sample is limited to those who were interviewed from ages 14 to 23.
* p<0.10, ** p<0.05, *** p<0.01

Table 22: Log-Wage of Working Individuals at Age 23

	Estimates	Estimates
Experience	0.061*	0.057*
	(0.03)	(0.03)
Experience Squared	-0.004	-0.004
	(0.00)	(0.00)
Education	0.043***	0.037***
	(0.01)	(0.01)
Currently Enrolled in School	-0.085*	-0.082*
	(0.05)	(0.05)
Cigarette Stock	-0.013	-0.010
	(0.01)	(0.01)
Alcohol Use Stock	0.039***	0.040***
	(0.01)	(0.01)
Marijuana Use Stock	-0.016	-0.011
	(0.01)	(0.01)
Total Arrests		-0.059***
		(0.02)
N	484	484

Data comes from the NLSY97. The sample is limited to those who were interviewed from ages 14 to 23.
* p<0.10, ** p<0.05, *** p<0.01

not a person works. Individuals using high level of substances may either be getting low wage offers because of their use or may just not like working. My structural model is able to control for selection, which allows me to better estimate the effect of substance use on wages.

B Functional Form of Utility Function

In this section I present the functional forms of the components of the utility function. In the following equations, $\mathbb{1}(\cdot)$ is an indicator that equals one if the argument is true. I assume that the deterministic part of the log-wage equation

$$
\begin{aligned}
\theta_w(sub_{it}, s_{it}) = \; & (\theta_1^w)^{\mathbb{1}(h_{it}=1)} (\theta_2^w + \theta_3^w Years\ of\ High\ School_{it} \\
& + \theta_4^w Years\ of\ College_{it} + \theta_5^w H_{it} + \theta_6^w H_{it}^2 + \theta_7^w \tau_{it} \\
& + \theta_8^w R_{it} + \theta_9^w unempr_{it} \\
& + \theta_{10}^w \mathbb{1}(cigs_{it} = 2) + \theta_{11}^w \mathbb{1}(cigs_{it} = 3) \\
& + \theta_{12}^w \mathbb{1}(alc_{it} = 2) + \theta_{13}^w \mathbb{1}(alc_{it} = 3) \\
& + \theta_{14}^w \mathbb{1}(mj_{it} = 2) + \theta_{15}^w \mathbb{1}(mj_{it} = 3)).
\end{aligned}
$$

I assume that the nonpecuniary utility of employment

$$
\begin{aligned}
\theta_n(sub_{it}, s_{it}) = \; & (\theta_1^n)^{\mathbb{1}(h_{it}=1)} (\theta_2^n + \theta_3^n H_{it} + \theta_4^n G_{it} + \theta_5^n \tau_{it} \\
& + \theta_6^n R_{it} + \theta_7^n unempr_{it} \\
& + \theta_8^n \mathbb{1}(h_{i,t-1} = 1) + \theta_9^n \mathbb{1}(h_{i,t-1} = 2) \\
& + \theta_{10}^n \mathbb{1}(cigs_{it} = 2) + \theta_{11}^n \mathbb{1}(cigs_{it} = 3) \\
& + \theta_{12}^n \mathbb{1}(alc_{it} = 2) + \theta_{13}^n \mathbb{1}(alc_{it} = 3) \\
& + \theta_{14}^n \mathbb{1}(mj_{it} = 2) + \theta_{15}^n \mathbb{1}(mj_{it} = 3)).
\end{aligned}
$$

I assume that the nonpecuniary utility of enrolling in school

$$
\begin{aligned}
\kappa_S\left(hc_{it}, sub_{it}, s_{it}\right) &= \kappa_1 + \kappa_2 Years\ of\ High\ School_{it} + \kappa_3 Years\ of\ College_{it} \\
&+\kappa_3 \mathbb{1}\left(enroll_{i,t-1} = 1\right) + \kappa_4 \mathbb{1}\left(g_{i,t-1} = 1\right) \\
&+\kappa_5 \mathbb{1}\left(h_{it} = 1\right) + \kappa_6 \mathbb{1}\left(h_{it} = 2\right) + \kappa_7 R_{it} \\
&+\kappa_8 \mathbb{1}\left(cigs_{it} = 2\right) + \kappa_9 \mathbb{1}\left(cigs_{it} = 3\right) \\
&+\kappa_{10} \mathbb{1}\left(alc_{it} = 2\right) + \kappa_{11} \mathbb{1}\left(alc_{it} = 3\right) \\
&+\kappa_{12} \mathbb{1}\left(mj_{it} = 2\right) + \kappa_{13} \mathbb{1}\left(mj_{it} = 3\right).
\end{aligned}
$$

Lastly, I assume that the utility gained from using substances

$$\begin{aligned}
\alpha\left(sub_{it}, s_{it}\right) &= \left(\alpha_{1}+\alpha_{2}\mathbb{1}\left(\mathscr{C}_{it}=0\right)+\alpha_{3}\mathscr{C}_{it}+\alpha_{4}\mathscr{A}_{it}+\alpha_{5}\mathscr{M}_{it}\right)\mathbb{1}\left(cigs_{it}=2\right)\\
&+\left(\alpha_{6}+\alpha_{7}\mathbb{1}\left(\mathscr{C}_{it}=0\right)+\alpha_{8}\mathscr{C}_{it}+\alpha_{9}\mathscr{A}_{it}+\alpha_{10}\mathscr{M}_{it}\right)\mathbb{1}\left(cigs_{it}=3\right)\\
&+\left(\alpha_{11}+\alpha_{12}\mathbb{1}\left(\mathscr{A}_{it}=0\right)+\alpha_{13}\mathscr{C}_{it}+\alpha_{14}\mathscr{A}_{it}+\alpha_{15}\mathscr{M}_{it}\right)\mathbb{1}\left(alc_{it}=2\right)\\
&+\left(\alpha_{16}+\alpha_{17}\mathbb{1}\left(\mathscr{A}_{it}=0\right)+\alpha_{18}\mathscr{C}_{it}+\alpha_{19}\mathscr{A}_{it}+\alpha_{20}\mathscr{M}_{it}\right)\mathbb{1}\left(alc_{it}=3\right)\\
&+\left(\alpha_{21}+\alpha_{22}\mathbb{1}\left(\mathscr{M}_{it}=0\right)+\alpha_{23}\mathscr{C}_{it}+\alpha_{24}\mathscr{A}_{it}+\alpha_{25}\mathscr{M}_{it}\right)\mathbb{1}\left(mj_{it}=2\right)\\
&+\left(\alpha_{26}+\alpha_{27}\mathbb{1}\left(\mathscr{M}_{it}=0\right)+\alpha_{28}\mathscr{C}_{it}+\alpha_{29}\mathscr{A}_{it}+\alpha_{30}\mathscr{M}_{it}\right)\mathbb{1}\left(mj_{it}=3\right)\\
&+\alpha_{31}age_{it}\mathbb{1}\left(cigs_{it}=2\right)+\alpha_{32}age_{it}\mathbb{1}\left(cigs_{it}=3\right)\\
&+\alpha_{33}age_{it}\mathbb{1}\left(alc_{it}=2\right)+\alpha_{34}age_{it}\mathbb{1}\left(alc_{it}=3\right)\\
&+\alpha_{35}age_{it}\mathbb{1}\left(mj_{it}=2\right)+\alpha_{36}age_{it}\mathbb{1}\left(mj_{it}=3\right)\\
&+\alpha_{37}\mathbb{1}\left(cigs_{it}\geq 2\right)\mathbb{1}\left(alc_{it}\geq 2\right)\\
&+\alpha_{38}\mathbb{1}\left(cigs_{it}\geq 2\right)\mathbb{1}\left(mj_{it}\geq 2\right)\\
&+\alpha_{39}\mathbb{1}\left(alc_{it}\geq 2\right)\mathbb{1}\left(mj_{it}\geq 2\right).
\end{aligned}$$

C Posterior Distribution

The posterior distribution is made up of the likelihood function and the prior distribution. For each individual i, I observe in the data a vector of optimal substance use and human capital accumulation choices $\left\{sub_{it}^{obs}, hc_{it}^{obs}\right\}_{t=1}^{T_i}$, a wage when an individual works, arrests, and grade completion status if the individual is enrolled in school. The additional latent variables that I simulate in order to augment the data are time persistent unobserved heterogeneity in preferences and skills μ_i; worker-firm match values $\psi_i\left(\cdot\right) = \left\{\psi_{ik}\right\}_{k=1}^{K_i}$; unobserved wage draws $w_i\left(\cdot\right) = \left\{\left\{w_{it}\left(hc_{it}\right)\right\}_{hc_{it}\neq hc_{it}^{obs}}\right\}_{t=1}^{T_i}$; unobserved parental transfers $\tilde{w}_i\left(\cdot\right) = \left\{\left\{\tilde{w}_{it}\left(hc_{it}\right)\right\}_{d_{it}\in D_{it}\left(s_{it}\right)}\right\}_{t=1}^{T_i}$, the latent variables that describe the probabilities of grade completion $G_i^*\left(\cdot\right) = \left\{G_{it}^*\right\}_{t=1}^{T_i}$ and arrest $R_i^*\left(\cdot\right) = \left\{R_{it}^*\right\}_{t=1}^{T_i}$; and the shocks to the substance use choice $\epsilon_i^{sub}\left(\cdot\right) = \left\{\epsilon_{it}^{sub}\right\}_{t=1}^{T_i}$. I denote the latent variables that I simulate as θ^*. I denote the parameters to be estimated as θ. The likelihood function represents the probability of observing the augmented data conditional on the parameters. I assume diffuse priors on all of the parameters. I denote the density of μ_i, ψ_i, ϵ_{it}^{sub}, and $\tilde{w}_{it}\left(hc_{it}\right)$ as $g\left(\cdot|\theta\right)$, where $g\left(\cdot|\theta\right)$ is the normal pdf given the distributional parameters in θ. Given this information, the joint posterior distribution of the parameters and the latent variables is proportional to

$$\Pr\left(\theta; \left\{\mu_i, \psi_i\left(\cdot\right), w_i\left(\cdot\right), \tilde{w}_i\left(\cdot\right), G_i^*, R_i^*, \epsilon_i^c\left(\cdot\right)\right\}_{i=1}^N\right) \propto \tag{19}$$

$$\prod_{i=1}^{N}\left[g\left(\mu_i|\theta\right) g\left(\psi_i|\theta\right)\right. \tag{20}$$

$$\cdot \prod_{t=1}^{T_i}\left[\mathbb{1}\left(R_{i,t}^* > 0\; if\; r_{it} = 1, < 0\; if\; r_{it} = 0\right) \cdot \Pr\left(R_{i,t}^*|s_{it}, sub_{it}^{obs}, \mu_i; \eta^a\right)\right. \tag{21}$$

$$\cdot \left(\mathbb{1}\left(G_{i,t}^* > 0\; if\; g_{it} = 1, < 0\; if\; g_{it} = 0\right)\right. \tag{22}$$

$$\left.\cdot \Pr\left(G_{i,t}^*|s_{it}, sub_{it}^{obs}, \mu_i; \eta^s\right)\right)^{\mathbb{1}(enroll_{it}=1)}$$

$$\cdot\; \mathbb{1}\left(\epsilon_{it}^{sub}\; is\; in\; bounds\right) \cdot g\left(\epsilon_{it}^{sub}|\theta\right) \tag{23}$$

$$\cdot \Pr\left(v_{it}\left(sub_{it}^{obs}, hc_{it}^{obs}|s_{it}, \theta^*; \theta\right) \geq v_{it}\left(sub_{it}\left(j\right), j|s_{it}, \theta^*; \theta\right)\; \forall\; j \in D_{it}\left(s_{it}\right)\right) \tag{24}$$

$$\cdot \prod_{j=1}^{D_{it}(s_{it})}\left[\mathbb{1}\left(\tilde{w}_{it}\left(j\right)\;\; satisfies\; CONS_{it} \geq C_{min}\right) \cdot g\left(\tilde{w}_{it}\left(j\right)|\theta\right)\right. \tag{25}$$

$$\cdot\;\; \Pr\left(\ln\left(w_{it}\left(j\right)\right)|s_{it}, \mu_i, \psi_i, sub_{it}\left(j\right); \theta\right)]]] \tag{26}$$

In the above equation, $\mathbb{1}\left(\cdot\right)$ is an indicator that equals one if the argument is true. The purpose of these terms is to make sure that the latent variables I simulate agree with the data. Line 21 guarantees that $R_{i,t+1}^*$ is greater than zero if an individual got arrested and is less than zero if he did not. Similarly the indicator in line 22 guarantees that $G_{i,t+1}^*$ is greater than zero if an individual completes a grade and less than zero if he did. Line 23 guarantees that ϵ_{it}^{sub} results in the observed substance use choice being the optimal continuous choice at the observed human capital accumulation choice. Lastly, line 25 guarantees that transfers are large enough so that consumption is larger than the minimum consumption level. As for the other terms in Equation 19,

$$\Pr\left(R_{i,t+1}^*|s_{it}, sub_{it}^{obs}, \mu_i^{arrest}; \eta_{arrest}\right) \;\; = \;\; \phi\left(R_{i,t+1}^* - \eta_{arrest}\left(\mu_i^{arrest}, sub_{it}^{obs}, s_{it}\right)\right)$$

and

$$\Pr\left(G^{*}_{i,t+1}|s_{it}, sub^{obs}_{it}, \mu^{pass}_{i}; \eta_{school}\right) = \phi\left(G^{*}_{i,t+1} - \eta_{school}\left(\mu^{arrest}_{i}, sub^{obs}_{it}, s_{it}\right)\right),$$

where ϕ is the standard normal pdf.

Let

$$\tilde{v}_{it}\left(sub_{it}\left(j\right), j|s_{it}, \mu_{i}, w_{it}, \tilde{w}_{it}, \epsilon^{sub}_{it}; \theta\right) = v_{it}\left(sub_{it}\left(j\right), j|s_{it}, \mu_{i}, w_{it}, \tilde{w}_{it}, \epsilon^{sub}_{it}; \theta\right) - \epsilon^{hc}_{it}$$

Given this,

$$\begin{aligned} &\Pr\Big(v_{it}\left(sub^{obs}_{it}, hc^{obs}_{it}|s_{it}, \mu_{i}, w_{it}, \tilde{w}_{it}, \epsilon^{sub}_{it}; \theta\right) \\ &\qquad \geq v_{it}\left(sub_{it}\left(j\right), j|s_{it}, \mu_{i}, w_{it}, \tilde{w}_{it}, \epsilon^{sub}_{it}; \theta\right)\ \forall\, j \in D_{it}\left(s_{it}\right)\Big) \\ &= \frac{\exp\left(v_{it}\left(sub^{obs}_{it}, hc^{obs}_{it}|s_{it}, \mu_{i}, w_{it}, \tilde{w}_{it}, \epsilon^{sub}_{it}; \theta\right)\right)}{\sum_{j\in D_{it}(s_{it})} \exp v_{it}\left(sub^{obs}_{it}, hc^{obs}_{it}|s_{it}, \mu_{i}, w_{it}, \tilde{w}_{it}, \epsilon^{sub}_{it}; \theta\right)}. \end{aligned}$$

Lastly, remember that today's wage offer is a function of today's substance use, which varies with the discrete choice. Therefore,

$$\ln\left(w_{it}\left(j\right)|s_{it}, \mu_{i}, \psi_{i}, sub_{it}\left(j\right); \theta\right) \sim N\left(\theta^{h}_{w}\left(sub_{it}\left(j\right), s_{it}\right) + \psi_{ik} + \mu^{w}_{i}, \sigma^{2}_{h}\right).$$

So,

$$\Pr\left(\ln\left(w_{it}\left(j\right)\right)|s_{it},\mu_i,\psi_i,sub_{it}\left(j\right);\theta_w\right)\propto$$
$$\sigma_h^{-1}\exp\left(-\frac{1}{2}\sigma_h\left(w_{it}\left(j\right)-\mu_{it}^h\left(j\right)\right)^2\right).$$

I divide the parameters and latent variables into blocks and draw from each block using the Metropolis-Hastings algorithm.