Marijuana Consumption and Education: Evidence from the NLSY97 and NSDUH

Golnaz Davalloo

A Thesis
in
The Department
of
Economics

Presented in Partial Fulfillment of The Requirements
For The Degree of
Doctor of Philosophy(Economics)
Concordia University
Montréal, Québec, Canada

December 2022

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: **Golnaz Davalloo**

Entitled: **Marijuana Consumption and Education:**

Evidence from the NLSY97 and NSDUH

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Economics)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the Final Examining Committee:

________________________ Chair
Dr. Christine Jourdan

________________________ External Examiner
Dr. Hai Van Nguyen

________________________ External Examiner to Program
Dr. Michel Magnan

________________________ Examiner
Dr. Ian Irvine

________________________ Examiner
Dr. Damba Lkhagvasuren

________________________ Supervisor
Dr. Jorgen Hansen

Approved by ________________________
Dr. Christian Sigouin, Graduate Program Director
Department of Economics

November, 04, 2022 ________________________
Dr. Pascale Sicotte, Dean
Faculty of Arts and Science

Abstract

Marijuana Consumption and Education:
Evidence from the NLSY97 and NSDUH

Golnaz Davalloo, Ph.D.
Concordia University, 2022

This dissertation analyzes the relationship between education and marijuana consumption among adolescents. Its three main objectives are: (i) analyze the effect of drug prevention programs on marijuana use; (ii) estimate the effect of marijuana use on educational attainment; and (iii) assess the roles of addiction, time-invariant unobserved heterogeneity, and time-dependent, time-varying shocks on persistent marijuana use.

To estimate these effects, I utilize micro-level data drawn from 15 waves of the National Longitudinal Study of Youth 1997 (NLSY97), covering the period from 1997 to 2011, and 13 waves of the National Survey on Drug Use and Health (NSDUH), covering the period from 2002 to 2014.

In chapter one, I create a pseudo-panel from repeated cross-sections of NSDUH and use the information on school-provided drug prevention programs. This information is not available in the NLSY97. I validate the pseudo-panel by comparing its main features with the corresponding ones for NLSY97. The results suggest that school-provided drug education decreases marijuana use, mainly by improving students' perception of the risks associated with marijuana use among adolescents.

Chapter two, which is co-authored with Jorgen Hansen, analyzes transitions into marijuana consumption jointly with grade transitions using data from the NLSY97. We allow for correlated unobserved heterogeneity that impacts both transitions within a discrete-time hazard framework. We estimate the impacts at different grade levels and find that they vary significantly. Average marginal effects indicate that using marijuana reduces next year's grade transition by 9.6 percentage points in high school

and 2.3 percentage points while in college. Adverse effects are more severe for male youth and students from disadvantaged family backgrounds.

The third chapter of my thesis, co-authored with Jorgen Hansen, analyzes persistence in marijuana consumption based on data from the NLSY97. We allow for three sources of persistence: pure state dependence (or addition), time-invariant unobserved heterogeneity, and persistent, idiosyncratic, time-varying shocks. We estimate a dynamic ordered Probit model using simulated Maximum Likelihood utilizing the intensity of consumption based on the number of days consumed per month. The results demonstrate a causal effect of previous use. In addition, the state dependence is significantly exaggerated when unobserved heterogeneity and serially correlated shocks are ignored.

To my parents

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Dr. Jorgen Hansen, for his help, support, patience, and insightful comments throughout my Ph.D. research and thesis preparation. His wisdom and depth of knowledge have always been a constant source of motivation and encouragement for me. Without his guidance, feedback, and assistance, I would not have been able to complete this Ph.D.

Besides my supervisor, I would like to thank my thesis defence chair Dr. Christine Jourdan and express my appreciation to the rest of my thesis defence committee, Dr. Hai Van Nguyen, Dr. Michel Magnan, Dr. Ian Irvine, and Dr. Damba Lkhagvasuren, for their precious time.

I would also like to acknowledge the former and current Graduate Program Directors, Dr. Effrosyni Diamantoudi, Dr. Szilvia Pápai, Dr. Damba Lkhagvasuren, and Dr. Christian Sigouin (current), and to the former and current Chairs, Dr. Greg LeBlanc and Dr. Jorgen Hansen. They were always available to support my journey in the Ph.D. program.

My special thank goes to the staff members of the Department of Economics, Ms. Elise Melancon, Ms. Lucy Gilson, Ms. Lise Gosselin, Ms. Bonnie Janicki, Ms. Mary-Ann Jirjis, Ms.Sandra Topisirovic, Ms. Melissa Faisal, and Ms. Emilie Martel, and Ms. Kelly Routly for their diligence, patience and help.

I am grateful for the unconditional love, encouragement, and support I have received from my mother, Parvin, and my father, Mohammad Mehdi since I can remember. In addition, I would like to thank my sister, Gordia and my brother-in-law, Amirali Amirhamzeh, for their support at this crucial time in my life, which made me feel at peace during the highs and lows. Lastly, I am grateful for the support provided by my dear friends Farnoosh Tajdivand, Hamed, and Farokh Darvish.

The fact that I am here would not be possible without all of them.

Contents

List of Figures

List of Tables

Introduction

Marijuana (cannabis)[1] has long been considered the most common gateway to hard drugs in the United States (Keyes et al., 2019; Sen et al., 2002), and its abuse is one of the substances associated with approximately 25% of all deaths in the United States annually (Boruch et al., 2003). Individuals who misuse drugs or alcohol, especially when those activities are severe and have problematic consequences, experience difficulties for themselves, their families, the workforce, and society as a whole (Tolan et al., 2007).

Until the 1980s, THC content in marijuana was less than 2% . In the 1990s, this ratio reached 4%, and between 1995 and 2015, it grew to 12%. A report from 2017 indicates that some popular strains of marijuana named "Girl Scout Cookie" in Colorado have THC contents ranging from 17 to 28% (Stuyt, 2018). This historical change in the level of THC increases the priority of studies regarding reasons of marijuana use and introducing the harmful health and social effect of this drug, as the problem goes beyond smoking weed with 2% THC mixed with tobacco as a joint in 1980s.[2]

The increase of THC levels in marijuana leads to a higher possibility of addiction and mental health disorder among adolescents (Di Forti et al., 2019). Broad psychological studies demonstrate that marijuana abuse harms young adolescents' brains (Brook et al., 2008; Cobb-Clark et al., 2013; Gonzalez and Swanson, 2012; Grant et al., 2010; Homel et al., 2014; Poudel and Gautam, 2017; Roebuck et al., 2004). Therefore, it is not surprising that adolescent substance abuse is one of the most serious public health issues in the United States. In spite of the legalization of recreational marijuana

[1] Marijuana and cannabis are used interchangeably throughout my thesis

[2] Despite the fact that marijuana and weed both refer to Cannabis, weed is defined as a "tobacco product" due to the way it is used.

use in nineteen states and the district of Columbia, which has altered the structure of marijuana consumption among adults in the United States, it remains illegal for adolescents. Therefore, it is important to prevent the popularity of marijuana among young people.

Numerous studies in different disciplines investigated the reasons for marijuana use and its effect. Some of these studies focused on personality traits (Pearson et al., 2018; Comeau et al., 2001), others on demographics such as family structure, socioeconomic situational, neuropsychological factors on marijuana consumption (Ewing et al., 2015), the effect of race on high-risk health issues, and social behaviour such as being primary substance abuse by youth offenders (Ewing et al., 2011). However, the primary focus of this study is on adolescents between the ages of 12 and 17 to determine whether school-based drug prevention programs influence marijuana use, whether marijuana use affects educational attainment, and why marijuana use persists.

In this regard, I use two famous American databases in the drug study, the National Longitudinal Survey of Youth - 1997 Cohort (NLSY97) and the National Survey on Drug Use and Health (NSDUH). To conduct the empirical analysis, I apply a number of techniques, including pseudo-panel data, hazard (survival) models, Polya models, transition probabilities, principal component factor analysis, and weighted fixed-effects linear models.

This study is organized into three separate chapters. Following the introduction, chapter one examines the impact of drug education on marijuana consumption using NSDUH data.

The primary objective of this chapter is to create a pseudo-panel from repeated NSDUH cross-sections in order to analyze the dynamic effects of drug prevention programs on marijuana use. School-based prevention programs have been added to many school curriculums over the past half-century. The programs differ across states and in terms of format and contents, and the main purpose is to inform students about the adverse effects of drug use and reduce risky behaviour among adolescents. Substantial research in different fields has been devoted to analyzing the effectiveness of these programs. Some results suggest that they are successful, while other results indicate the opposite. In this chapter, I examine school-based drug education's direct and indirect effects on marijuana use at the national level.

The results show that school-based drug education reduces marijuana use through the indirect channel of risk perception. However, the direct effect of drug education on marijuana use is limited. In addition, I show that the age of receiving a school-based drug prevention program is critical. According to my results, this type of drug education should be offered before age 15 for having the highest level of effectiveness on marijuana use.

Chapter two examines the impact of marijuana use on educational attainment using data from NLSY97. Specifically, I analyze the transition into marijuana use jointly with grade transitions, allowing for the existence of correlated unobserved heterogeneity. A bivariate survival model is estimated to determine the impact of marijuana use on the continuation of education at different grade levels. According to the results, marijuana use reduces the probability of staying in high school by 9.6 percentage points and by 2.3 percentage points while in college. In addition, estimates show that marijuana use affects males' educational attainments more than females, as well as individuals with weaker family backgrounds and Hispanic students.

Marijuana consumption and its associated risks of abuse are heterogeneous in the population. Marijuana abuse may lead to negative or undesired consequences that are likely to vary with consumption patterns. Chapter three examines the dynamics of marijuana use and changes in consumption patterns over time. Utilizing a sample from the NLSY97, I estimate a dynamic ordered probit model using simulated maximum likelihood techniques. I allow for three sources of persistence in marijuana consumption: true state dependence or addiction, time-invariant unobserved heterogeneity and persistence in idiosyncratic, time-varying shocks.

In this chapter, the consumption patterns are based on the number of days per month individuals use marijuana over time. I group consumption into three categories:

none, moderate and heavy, where moderate use corresponds to consumption less than 9 times per month, and heavy use corresponds to the consumption of 10 days or more. The results indicate that the probability of consuming moderate levels of marijuana in year t is 4.6 percentage points higher if the person consumed the same level of marijuana in year $t-1$, relative to not using any marijuana in year $t-1$. This constitutes a relative effect that is close to 50 percent, given the observed proportions of moderate consumption that are observed in the data. Further, the probability of consuming heavy levels of marijuana in year t is 4.3 percentage points higher if the person consumed the same level of marijuana in year $t-1$, relative to not using any marijuana in year $t-1$.

After separating persistence into three components, I find that persistent unobserved heterogeneity plays a significant role in the persistence of both heavy marijuana consumption and moderate use, accounting for 40% and 32% of the overall persistence, respectively. Moreover, the true or causal state dependence is greater for moderate consumption than heavy consumption by 47% and 33%, respectively, and the remainder is due to persistence in time-varying random shocks.

Chapter 1

School-Based Drug Prevention Programs and Marijuana Use

Abstract

I examine the effects of school-based prevention programs on marijuana consumption among adolescents by using data from the National Survey on Drug Use and Health (NSDUH), a series of cross-sectional surveys from 2002 to 2014. In spite of the fact that it is impossible to follow individuals over time at NSDUH, this study provides a pseudo-panel methodology to examine the effects of school-based drug education on marijuana use among different groups of individuals over time.

According to my research, a school-based drug prevention program can effectively reduce marijuana consumption. However, its effect is achieved through indirect channels influencing students' perceptions of marijuana's risk. For maximum effectiveness, it is necessary to deliver school-based drug education to adolescents before age 15 when they have not yet been exposed to marijuana by their peers.

JEL Code:I23,D10, D91

Keywords: Marijuana use; school-based drug prevention programs, reported risk perception, family background; peers who use marijuana;

1.1 Introduction

The current study examines the effects of school-based drug prevention programs on reducing marijuana use among adolescents in direct and indirect ways. The use of illicit drugs can generally be curbed through prevention, treatment, and enforcement. Both treatment and enforcement were subjected to cost-effectiveness analyses in the mid-1990s. More recently, the researchers' focuses were shifted to the cost-effectiveness of prevention (Caulkins et al., 2002; Botvin and Griffin, 2006; Chatterji, 2006; Vigna-Taglianti et al., 2009; Gabrhelik et al., 2012; Čurová et al., 2021).

Botvin and Griffin (2006) reviewed research evaluating school-based prevention programs from 1964 to 2002. During the 1970s and 1980s, there were no evaluation components available, or the methodologies were unreliable since they assessed knowledge and attitudes rather than actual drug use. Several recent studies indicate that school-based drug prevention is ineffective in the long term (Gabrhelik et al., 2012; Orosová et al., 2020; Čurová et al., 2021), while others indicate that its effectiveness is limited (Vigna-Taglianti et al., 2009).

Around the world, there are various types of school-based prevention programs, but their common goal is to prevent and reduce drug use among adolescents. Several of these programs, such as health enhancement curriculum (HEC), emphasize nutrition, physical activity, and general health while indirectly addressing drug abuse. In contrast, other programs, such as school-community intervention programs (SCI), focus directly on drug abuse and have specific plans for school-based drug prevention program (Flay et al., 2004).

In this study, I collected adolescents' self-reported responses regarding direct school-based drug prevention programs. Therefore, the findings can be generalized to any school-based drug prevention program that includes direct drug education in its curriculum. Moreover, the current study examines the effects of a school-based drug prevention program on marijuana utilization behaviour. To accomplish this target, I created a pseudo-panel from the cross-sectional surveys of the National Survey on Drug

Use and Health (NSDUH)(2002-2014). It is unfortunate that the National Longitudinal Survey of Youth 1997 (NLSY97), a large panel data with similar information regarding marijuana use and family structure, cannot be used considering its lack of information regarding drug education at school. However, I use the NLSY97 as a reference for evaluating the pseudo-panel.

This study found that school-based drug prevention programs do not directly affect marijuana consumption, similar to findings in previous studies such as Shackleton et al. (2016) and Perry et al. (2003). However, it does not mean that school-based drug prevention programs don't work. Based on the analysis and discussion in the following sections, it appears that school-based drug education helps adolescents become more aware of the risks associated with using marijuana, allowing them to decrease their use of marijuana under certain conditions. First of all, the effects of school-based drug education do not last more than a year, and booster courses are necessary to maintain awareness levels. Second, the right timing of school-based drug education is crucial. In other words, when students have already been exposed to marijuana, the indirect effect of drug education at school diminishes. School-based drug education must also be delivered to students before age 15, as 15- and 16-year-olds tend to follow their peers and care less about the harm marijuana produces.

The following is an outline of the remaining sections of this chapter. First, the following section summarizes the literature review concerning a school-based drug prevention program, its effects on marijuana use, and pseudo-panel data. Next, a description of the data is presented in section 3, followed by an explanation of the economic modelling procedure in section 4. Finally, in section 5, the results are discussed, and the chapter's conclusion is in section 6.

1.2 Literature

1.2.1 The Effect of School-Based Drug Prevention Programs on Marijuana Use

In North America and Europe, school-based drug prevention programs (SBDPs) have been extensively studied and examined from a variety of angles. Several studies have assessed the effectiveness of these programs according to their target behaviours such as smoking, drug abuse, alcohol abuse, and risky behaviour in general (Shackleton et al., 2016; Wilson et al., 2001; Botvin et al., 1990a). Some researchers analyzed the efficiency of programs based on their categories such as comparing interactive and non-interactive prevention programs, direct or indirect school-based drug prevention programs, and long-term or short-term programs (Tanner-Smith et al., 2018; Soole et al., 2008; Fletcher et al., 2008; Rosenbaum, 2007; Bond et al., 2004; Tobler et al., 1999; White and Pitts, 1998; Botvin et al., 1995, 1990b; Hansen, 1992). Finally, other researchers analyzed the success of these programs based on different sample categories such as boys versus girls and school grade levels (Onrust et al., 2016; Steinberg, 2010; Birkeland et al., 2005; Perry et al., 2003; Wilson et al., 2001)

Wilson et al. (2001) conducted a meta-analysis for studies from 1990 to 2001 regarding the effect of school-based prevention on four categories of crime, substance use, dropping out of school, and other behavioural problems. The results show the significant power of drug prevention programs in controlling substance abuse at schools (Botvin et al., 1995; Gerstein and Green, 1993; Hansen, 1992; Botvin et al., 1990b; Dryfoos, 1990). More recently, Shackleton et al. (2016) evaluated 7,544 unique references and 22 reviews about school-based prevention programs that are published after 1980. Despite the many studies demonstrating the effectiveness of school-based prevention programs in promoting sexual health, preventing bullying, and preventing smoking, no high-quality studies have documented the effectiveness of multi-component interventions for reducing alcohol or drug use.

According to Fletcher et al. (2008), changing the social environment of schools in order to improve teacher-student relationships and foster positive school climates, in addition to enhancing adolescents' knowledge of refusal and developing negotiation skills, as well as modifying peer norms can reduce drug use and other risky health behaviours. These results are in line with those of Tobler et al. (1999) that believe interactive programs with intensive sessions are more effective than non-interactive ones because they engage families and peers. Furthermore, expanding the number of the participants by over 400, decreases the effectiveness of the programs.

According to Tanner-Smith et al. (2018)'s mega-analyses evaluation, prevention programs which are only based on the presentation of drug information, and do not provide interactive workshops or activities, rarely result in effective prevention.

The Gatehouse project, a school-based prevention program, is evaluated by Bond et al. (2004) and Hansen (1992) based on the implementation of four waves in two different periods of its performance among eighth-grade students in Australia. Accordingly, both studies find that the marijuana usage rate among young people in the intervention group decrease significantly compare to baseline in wave four. A few years later, Rosenbaum (2007) evaluate the Drug Abuse Resistance Education (D.A.R.E), another school-based prevention program. According to Rosenbaum's results, the program's beneficial effects on improving the students' knowledge of the drug, attitudes toward the police, and social skills would soon disappear within one or two years.

According to Bond et al. (2004) study, schools should implement prevention programs consisting of booster programs every year to significantly and sustainably reduce marijuana consumption. Therefore, intervention programs cannot be quick fixes or simple solutions. Those programs can only succeed if schools and communities commit to them for a long time.

In 2008, a meta-analysis was conducted by Soole et al. (2008) to assess the impacts of school-based drug prevention programs (SBDP) on illicit drug use. Similar to previous studies, this systematic review shows that successful programs are often characterized by high levels of interaction and occur most frequently during middle school. Furthermore, Soole et al. (2008) and White and Pitts (1998)'s short-term and long-term analyses suggest that school-based drug programs' effect on drug consumption diminishes over time. Additionally, Fletcher et al. (2008) is another systematic review that examined the impact of intervention programs on drug use in schools. According to the results, interactive prevention programs can improve a school's social climate and the relationship between teachers and students, resulting in fewer drugs being used than simply providing information on drugs.

The findings of Wilson et al. (2001) indicate that previous studies have not adequately assessed the effectiveness of school-based prevention programs. Nevertheless, the authors determine that all school-based prevention programs can have at least a small impact on decreasing alcohol and drug use and changing other behaviour problems. Moreover, the school-based prevention programs are not comparable because of their different characteristics.

Perry et al. (2003) compared a control group of students in seventh grade with a group of students who received two different intervention programs, D.A.R.E, implemented in 16 schools, and D.A.R.E plus, implemented in 8 schools in the United States for two years. It is evident from the results that one of these programs had a meaningful effect on reducing drug use among boys. However, neither of these programs had any effect on girls.

Onrust et al. (2016) conducted a systematic meta-analysis to examine the effectiveness of school-based prevention programs for four target groups:

1) Children (elementary school)
2) Early adolescents (grades 6 and 7)
3) Middle adolescents (grades 8 and 9)
4) Late adolescents (grades 10 to 12)

These four groups of individuals have exhibited a high level of heterogeneity in their responses to prevention programs. For example, in elementary school, developing essential skills such as social skills, self-control, and problem-solving abilities will significantly influence youth more than direct prevention initiatives. Thus, prevention programs are unsuitable for children in elementary school. The best candidates for prevention programs are early adolescents, as they will be exposed to drugs in real life, and having essential knowledge is crucial.

Middle adolescents tend to be compassionate about their peers' expectations, opinions, and needs during school, so programs emphasizing peer pressure are ineffective. Late adolescents need more preparation for adult life than early and middle adolescents. As a result, peer approval is not as crucial as social acceptance in developing a personal social identity. Unfortunately, these programs are not helpful to late adolescents with substance abuse issues since drug consumption has become part of their identity.

The findings from previous studies indicate that school-based prevention programs have varying levels of effectiveness, ranging from quite effective to almost ineffective in reducing marijuana use, delaying its onset, or preventing its use. However, these programs are still necessary as they provide many public benefits that significantly exceed their costs (Caulkins et al., 2002). Results in Tobler et al. (1999) indicate that insufficient analyses regarding drug prevention programs can be the reason of believing these programs are not effective. Also, school-based prevention programs have evolved in their contents and delivery methods over the past four decades, making them more successful than in the past. Furthermore, the cost and preliminary studies are not the only reasons to do so. Developing and maintaining prevention programs in schools is far more manageable than addressing treatment and enforcement. Additionally, even modest impacts on marijuana use have been shown to influence educational attainment (Hansen and Davalloo, 2022), which is more easily addressed by appropriate policies than by trying to improve family circumstances or changing peers to access the same level of educational attainment (Chatterji, 2006; Hansen and Davalloo, 2022).

Accordingly, these studies have provided some evidence that may contribute to the success of school-based prevention programs. The following factors have been identified as necessary for a successful school-based drug prevention program:

1) Teachers and parents participate in the program,

2) Program engages direct school-based drug prevention in its curriculum,

3) Long-term program with sustainable and booster courses, and

4) Program begins no later than grades six and seven, since the majority of students have not been exposed to drugs yet.

In contrast to previous studies, the current study analyzes the effect of direct drug education at school on marijuana use without restricting the study to specific school-based drug education programs, schools, states or groups.

1.3 Data

Throughout this chapter, I analyze data from the National Survey on Drug Use and Health (NSDUH), an annual series of cross-sectional surveys conducted by the federal government since 1971. These surveys collect information from civilians and non-institutionalized adults aged 12 and older every year. Due to the nature of these cross-sectional surveys, it is impossible to capture unobserved individual characteristics and follow up with participants regarding the effect of specific actions in the future. Therefore, to examine the effectiveness of school-based drug prevention programs on marijuana use, I constructed pseudo-panel data by classifying individuals with similar characteristics into similar cohorts.

1.3.1 Pseudo-Panel Data

A panel data preparation process can be costly, leading to high loss rates over an extended period of time as attrition may cause a selection bias. In contrast, yearly independent cross-sectional data are relatively easy to access. They could act as panel data if arranged as pseudo-panels when actual panel data is not available (Bernard et al., 2011). Moreover, according to Nijman and Verbeek (1990) and Deaton (1985), pseudo-panels have 30 to 70 percent cost advantages over panel data with similar observations quantity, and they can run for long periods of time without the risk of attrition bias.

The pseudo-panel techniques allow using cross-sectional information regarding categorized individuals over time, as suggested by Deaton (1985). In this way, pseudo-panels represent stable groups of individuals rather than individuals over time, and the mean of observed individual variables inside each cohort is replaced in the panel as the cohort variable value. That is a linear transformation from the individual level to the cohort level. As a result, the individual fixed effect is substituted by its pseudo-panel data counterpart. Therefore, the cohort fixed effect can be estimated as corresponding to the individual fixed effect level (Guillerm, 2017).

Deaton (1985) was among the first to propose considering each cohort's average data as an observation and introduced a cohort as a group of individuals with similar characteristics that remain constant in the surveys over time. The cohorts' observations are the average of categorized individuals' data, not the whole population. As a result, measurement error is a common problem and Deaton (1985) introduced error-in-variables techniques for this issue.

The trade-off between panel data and pseudo-panel is the costly process of precise data collection subject to attrition compared to comprehensive data subject to measurement errors. However, increasing the number of individuals in each cohort decreases the measurement errors, and in empirical applications, these errors are generally ignored (Blundell et al., 1994; Browning et al., 1985; Moffitt, 1993). For example, Bernard et al. (2011) conducted a pseudo-panel using Deaton's (1985) study with 25 cohorts and, on average, 131 households in each cohort and ignored measurement errors as each cohort consisted of a sufficiently large number of individuals.

As Verbeek and Nijman (1992) emphasizes, choosing characteristics that can categorize individuals into different cohorts is the first essential factor for a successful pseudo-panel. Furthermore, a second principal factor is the consideration of the arbitrage between the number of cohorts and the number of participants within each cohort. For example, suppose the number of participants in each cohort is small, but the number of cohorts is large. In this case, each cohort average may not be close to the population's average, and the accuracy of the pseudo-panel decreases.

According to Gardes et al. (2005), all time-invariant fixed effects may not be removed by first differencing or fixed effect estimation when using a pseudo-panel. It is primarily due to the fact that, even though individuals within each cohort share similar characteristics, each cohort consists of different individuals.

The current study is aware of the limitations introduced by Gardes et al. (2005). However, creating cohorts, as groups of 150 individuals in average with time invariant and distinguishable characteristics, decreases measurement error and the fixed-effects estimator is asymptotically consistent (Verbeek and Nijman, 1992; McKenzie, 2004; Baltagi et al., 2015; Guillerm, 2017). To estimate the causal effect of school-based drug education on marijuana use without peeking correlated effects, I create this pseudo-panel data that allow me to control on time-invariant unobserved heterogeneity and let me to look at its dynamics and persistence in its time lags.

1.3.2 Pseudo-Panel Data Creation

An adequate pseudo-panel must satisfy several criteria. First, individuals must be categorized by their constant characteristics over time. Second, characters of categorizing must be introduced so that each individual can only be classified into one cohort (Guillerm, 2017). Finally, the number of individuals in each cohort must be large enough to decrease the measurement error on intra-cohort variable means. Based on Verbeek and Nijman's study conducted in 1992, it is shown that the level of efficiency of the pseudo-panel is boosted to an optimal level by having on average 100 to 200 individuals in each cohort.

In this section analysis, I create a pseudo-panel data from the information of 722,653 individuals who participated in NSDUH from 2002 to 2014, using the following restrictions:

1) Individuals must be between 12 to 17 years old.[1]

[1] 492,386 individuals are older than 17.

2) Only individuals who answered questions regarding risk attitudes, will be considered in the pseudo-panel.[2]

3) Each specific cohort joins the panel at age 12 and stays until age 17.

4) In the first year of the panel (age 12), none of the cohorts have used marijuana yet.[3]

5) All individuals are enrolled in school.

6) Each cohort must stay in the panel at least for 2 periods.

After these restrictions, 168,209 individuals remain for the creation of the pseudo-panel.[4] Table 1.1 visualizes the combination of individuals at different ages and years who are used to create the pseudo-panel.

Specifically, I used three time-invariant characteristics that divided individuals into separate cohorts: Gender (male, female), race (Black, Hispanic, non-Black and non-Hispanic) and risk attitude (Never, seldom and sometimes/always like to do risky things). The result is a pseudo-panel with 18 cohorts in each age group yearly ($2\times3\times3$), see Table 1.2 below.

[2]Among 230,267 individuals ages 12 to 17, 0.8% (1,772 individuals) did not answer, "Do you like to do risky things?"

[3]With the assumption of not using marijuana before the age of 13, endogeneity and initial condition are not considered as problems for the dynamic model anymore.

[4]More information is available in Table B.3

Table 1.1: Individuals age-year categories in NSDUH cross-sectionals

Age Year	12	13	14	15	16	17	Total
2002	2,878	-	-	-	-	-	2,878
2003	2,745	2,787	-	-	-	-	5,532
2004	2,693	2,907	2,819	-	-	-	8,419
2005	2,731	2,821	2,881	2,930	-	-	11,363
2006	2,590	2,786	2,793	2,876	2,834	-	13,879
2007	2,539	2,651	2,603	2,778	2,789	2,583	15,943
2008	2,446	2,564	2,636	2,830	2,918	2,682	16,076
2009	2,373	2,519	2,675	2,863	2,808	2,743	15,981
2010	2,672	2,838	2,762	2,889	2,917	2,861	16,939
2011	2,768	2,979	2,936	3,017	3,169	2,970	17,839
2012	2,646	2,580	2,607	2,750	2,842	2,743	16,168
2013	2,573	2,747	2,817	2,820	2,843	2,705	16,505
2014	-	2,117	2,181	2,182	2,187	2,020	10,687
Total	31,654	32,296	29,710	27,935	25,307	21,307	168,209

Table 1.2 presents the 1,116 observations from 216 cohorts based on the combination of age year distribution. These cohorts are in the pseudo-panel between two and six periods. Information regarding marijuana consumption and school-based drug prevention programs among aged 12 adolescents are used to analyze marijuana consumption at age 13 (More information about details on the sample selection of NSDUH pseudo-panel is in Appendix B, Table B.4).

Table 1.2: Cohort age-year categories in pseudo-panel

Age Year	12	13	14	15	16	17	Total
2002	18	-	-	-	-	-	18
2003	18	18	-	-	-	-	36
2004	18	18	18	-	-	-	54
2005	18	18	18	18	-	-	72
2006	18	18	18	18	18	-	90
2007	18	18	18	18	18	18	108
2008	18	18	18	18	18	18	108
2009	18	18	18	18	18	18	108
2010	18	18	18	18	18	18	108
2011	18	18	18	18	18	18	108
2012	18	18	18	18	18	18	108
2013	18	18	18	18	18	18	108
2014	-	18	18	18	18	18	90
Total	216	216	198	180	162	144	1,116

Table 1.3 and Table 1.4 show the transformation of NSDUH Cross-sectionals data to NSDUH pseudo-panel by comparing the distribution of age and year in both databases. In Table 1.3 the percentages of age combinations for the NSDUH pseudo-panel and the NSDUH cross-sectional data, in column 3 and column 5, are very similar. Also, columns 3 and 5 in Table 1.4 are similar in terms of percentages of observations in each year. Therefore, NSDUH's demographic composition did not change as a result of pseudo-panel creation.

Table 1.3: Age categories comparison between NSDUH and pseudo-panel

Age	Pseudo-panel data		NSDUH	
	N	%	N	%
13	216	24	32,296	23.65
14	198	22	29,710	21.8
15	180	20	27,935	20.45
16	162	18	25,307	18.5
17	144	16	21,307	15.6
Total	900	100	136,555	100

By contrast, comparing the percentage distributions of cohorts by gender, race, and risk attitude in pseudo-panels and the NSDUH cross-sectional data indicates very different results. To put it differently, it is the creation of a panel with a proportionately equal mix of all 18 cohorts (see Table 1.5). Due to this equality, all cohorts have the same importance regardless of their proportion in the NSDUH. Using a weighted pseudo-panel in the analysis produces results that reflect the sample combination of the NSDUH cross-sections database.[5]

Table 1.4: Year categories comparison between NSDUH and pseudo-panel

Year	Pseudo-panel data		NSDUH	
	N	%	N	%
2003	18	2	2,787	2
2004	36	4	5,726	4.2
2005	54	6	8,632	6.3
2006	72	8	11,289	8.3
2007	90	10	13,404	9.8
2008	90	10	13,630	10
2009	90	10	13,608	10
2010	90	10	14,267	10.4
2011	90	10	15,071	11
2012	90	10	13,522	9.9
2013	90	10	13,932	10.2
2014	90	10	10,687	7.9
Total	900	100	136,555	100

Table 1.5: Race, gender, risk attitude sample ratio
pooled NSDUH and pooled pseudo-panel

	Gender	Male		Female	
Race	Risk	NSDUH	Pseudo-panel	NSDUH	Pseudo-panel
Non-Black, non-Hispanic	Never	0.0743	0.0556	0.1031	0.0556
	Seldom	0.1294	0.0556	0.1401	0.0556
	Sometimes & always	0.1494	0.0556	0.1030	0.0556
Black	Never	0.0271	0.0556	0.0303	0.0556
	Seldom	0.0183	0.0556	0.0206	0.0556
	Sometimes & always	0.0191	0.0556	0.0152	0.0556
Hispanic	Never	0.0279	0.0556	0.0321	0.0556
	Seldom	0.0262	0.0556	0.0283	0.0556
	Sometimes & always	0.0311	0.0556	0.0246	0.0556

[5]More detailed explanation will be provided upon request.

Table 1.6: Descriptive statistics of NSDUH weighted pseudo-panel (2002-2014)

	Observations	*Mean*	*Std.dev.*	*Min*	*Max*
Used marijuana(t)	900	0.124	0.112	0	0.477
Males	450	0.122	0.106	0	0.422
Females	450	0.127	0.117	0	0.477
Used marijuana$_{(t-1)}$	900	0.074	0.091	0	0.477
Males	450	0.071	0.084	0	0.358
Females	450	0.077	0.098	0	0.477
drank alcohol$_{(t-1)}$	900	0.203	0.168	0	0.772
Males	450	0.178	0.148	0	0.644
Females	450	0.228	0.183	0	0.772
Smoked cigarettes$_{(t-1)}$	900	0.082	0.079	0	0.417
Males	450	0.074	0.069	0	0.373
Females	450	0.091	0.088	0	0.417
School-based drug prevention program$_{(t-1)}$	900	0.891	0.312	0	1
Males	450	0.851	0.356	0	1
Females	450	.931	0.253	0	1
reported risk perception(t)	900	1.592	0.966	0	3
Males	450	1.424	0.958	0	3
Females	450	1.760	0.946	0	3
Caring family$_{(t-1)}$ **(more explanation in 1.12)**	900	0.025	0.294	-1.015	0.643
Males	450	0.058	0.255	-0.671	0.643
Females	450	-0.008	0.325	-1.015	0.617
Peers use marijuana(t)	900	1.056	0.724	0	3
Males	450	0.920	0.676	0	3
Females	450	1.191	0.746	0	3
Real family Income$\div 10,000$ **(t)**	900	3.876	1.038	2.040	6.19
Males	450	3.884	1.050	2.157	6.19
Females	450	3.868	1.027	2.040	5.97
Live in metropolitan city(t)	900	0.782	0.413	0	1
Males	450	0.791	0.407	0	1
Females	450	0.773	0.419	0	1
Age (t)	900	14.8	1.401	13	17
Males	450	14.8	1.402	13	17
Females	450	14.8	1.402	13	17

1.3.3 NSDUH Pseudo-Panel Data Validation

To validate the pseudo-panel, I used NLSY97 as reference panel data with the same assumptions I used for the pseudo-panel.

Table 1.7: Individuals age-year categories NLSY97 compatible with final pseudo-panel

Age Year	12	13	14	15	16	17	Total
1997	-	-	-	-	-	-	-
1998	-	968	-	-	-	-	968
1999	-	-	966	-	-	-	966
2000	-	-	-	956	-	-	956
2001	-	-	-	-	951	-	951
2002	-	-	-	-	-	948	948
Total	-	968	966	956	951	948	4,789

Table 1.8: Age distribution in the pseudo-panel and NLSY97 panel

Age	NSDUH Pseudo-panel		NLSY97 Panel	
	N	%	N	%
13	216	24	968	20.2
14	198	22	966	20.1
15	180	20	956	20
16	162	18	951	19.9
17	144	16	948	19.8
Total	900	100	4,789	100

Table 1.9 compares three databases pooled NSDUH, NSDUH pseudo-panel, and NLSY97. The first part of the table shows that the gender combinations in all three databases are similar. As noted in the second part of this table, the percentage of marijuana users in each age group is similar for the pooled NSDUH and NSDUH pseudo-panel. However, compared to the two other databases, the percentages of marijuana users in the NLSY97 database are higher.

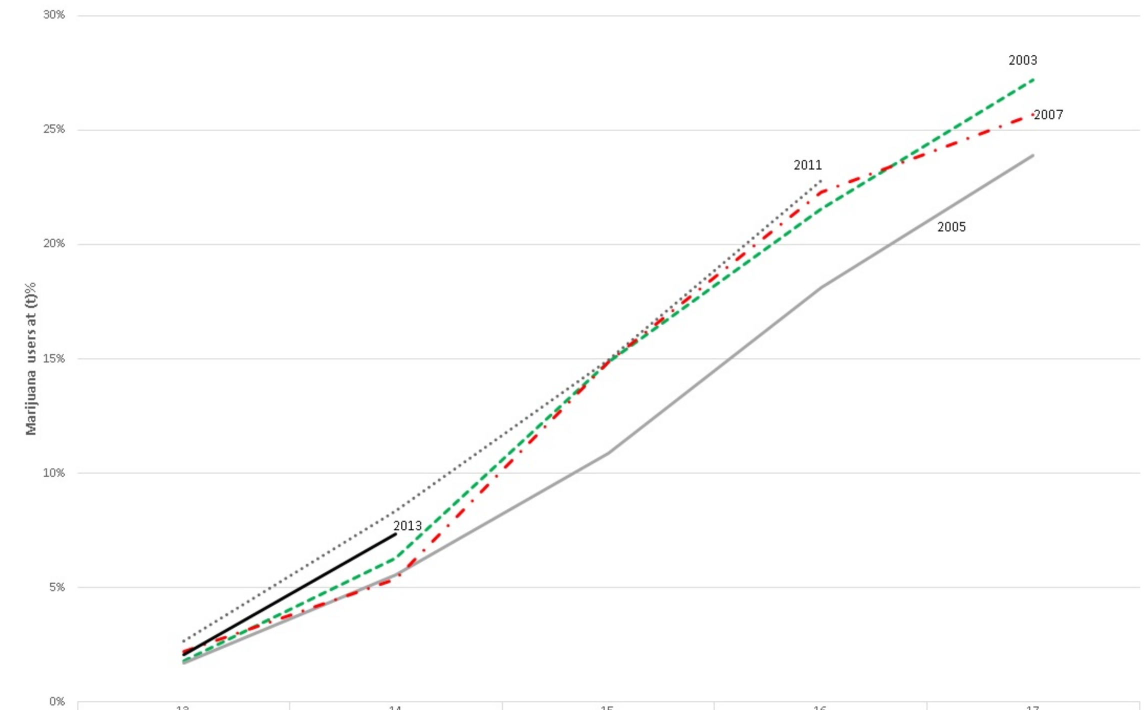
Table 1.9 shows differences due to these databases' different structures and timelines. First, based on our assumptions in this chapter, new 12-year-olds are added yearly to NSDUH. NLSY97, however, only includes 12-year-olds from 1997. The second factor to consider when comparing two databases is time. For example, the NLSY97 in 2002 reported that 30.8 percent of 17-year-olds had used marijuana, which is almost the same percentage as the NSDUH in 2002. However, this chapter did not consider the 17-year-olds who were participants in NSDUH 2002 (more details are in Table B.3 and

Table 1.2). Finally, among those who started marijuana before age 18, the percentage of the age of initiation at each age group shown in the third part of the table, is similar for all three databases.

Table 1.9: Comparing descriptive statistics of pooled NSDUH, NSDUH pseudo, and NLSY97

	NSDUH Pooled		NSDUH Pseudo-panel		NLSY97 Panel	
Characteristics	N	%	N	%	N	%
Number of observation	136,555		900		4,789	
Number of members	136,555		216		968	
Male	68,819	50.4%	108	50%	488	50.4%
Female	67,739	49.6%	108	50%	480	49.6%
Used marijuana at time(t) based on age level						
13 Overall		1.6%		2%		4.1%
Male		1.4%		1.6%		4.1%
Female		1.7%		2.4%		4.17%
14 Overall		6.1%		6.6%		19.3%
Male		5.9%		6.2%		18.72%
Female		6.3%		7.1%		19.79%
15 Overall		13.6%		13.7%		26.7%
Male		13.1%		13.1%		28.31%
Female		14.1%		14.3%		25%
16 Overall		20.3%		20.4%		31.3%
Male		20.7%		20.3%		32.22%
Female		19.9%		20.5%		30.43%
17 Overall		26.1%		25.6%		30.8%
Male		27.9%		26.4%		33.89%
Female		24.3%		24.9%		27.7%
Age of marijuana initiation, among who started before 18						
13		24%		19.9%		26%
14		54%		53.1%		47%
15		80%		81%		71%
16		96%		92.1%		88%
17		100%		100%		100%

Figure 1.1: Yearly comparison of the percentage of marijuana users in each age category in NSDUH pseudo-panel(2002-2014)



	13	14	15	16	17
2003	1.807%	6.281%	14.868%	21.573%	27.183%
2005	1.683%	5.543%	10.858%	18.110%	23.883%
2007	2.205%	5.389%	14.852%	22.275%	25.672%
2011	2.690%	8.370%	14.950%	22.783%	
2013	2.086%	7.340%			

1.4 Estimation

1.4.1 Fixed Effect Linear Pseudo-Panel Estimation

The basic idea of demonstrating a pseudo-panel starts with using a set of T independent cross-sections as below:

$$y_{it} = x'_{it}\beta + \mu_i + \upsilon_{it} \qquad t = 1, 2, \ldots, T \tag{1.1}$$

In equation 1.1, (it) refers to individual i at time t that is different from i at time t'. After categorizing individuals in cohorts with specific fixed characteristics throughout the T periods, and creating C sets of cohorts that contain averages of observations over individuals in each cohort. Equation 1.2 is extracted from equation 1.1.

$$\overline{y}_{ct} = \overline{x}'_{ct}\beta + \overline{\mu}_{ct} + \overline{\upsilon}_{ct} \qquad c = 1, 2, \ldots, C \quad ; \quad t = 1, 2, \ldots, T \tag{1.2}$$

where $\overline{y}_{ct} = \sum \frac{y_{i,t}}{n_{c_{(t)}t}} \forall i \epsilon c$ and $\overline{x}_{ct} = \sum \frac{x_{i,t}}{n_{c_{(t)}t}} \forall i \epsilon c$.

The time-invariant parameter μ_i, known as the individual fixed effect in equation 1.1, has an equivalent in equation 1.2 known as fixed cohort effect $\overline{\mu}_{ct}$ that varies over time. $\overline{\mu}_{ct}$ most likely correlates with the x_{it} and lead to inconsistent random effect estimation. However, with the assumption of $n_{c(t)} \to \infty$ for each cohort C at a fixed time (number of observations in each cohort be very large), it is expected that $\overline{\mu}_{ct}$=$\overline{\mu}_c$ and equation 1.2 changes to equation 1.3 (Baltagi et al., 2015).

$$\overline{y}_{ct} = \overline{x}'_{ct}\beta + \overline{\mu}_c + \overline{\upsilon}_{ct} \qquad c = 1, 2, \ldots, C \quad ; \quad t = 1, 2, \ldots, T \tag{1.3}$$

where $\overline{\upsilon}_{ct}$ varies with cohorts and time and can be thought as the usual disturbance or the remainder disturbance in the regression (Baltagi et al., 2015). The mean over time is

$$\overline{y}_c = \overline{x}'_c\beta + \overline{\mu}_c + \overline{\upsilon}_c. \tag{1.4}$$

Based on the within cohort transformation $\tilde{y}_{ct} = \overline{y}_{ct} - \overline{y}_c$ (subtracting equation 1.4 from equation 1.3)

$$\overline{y}_{ct} - \overline{y}_c = (\overline{x}'_{ct} - \overline{x}'_c)\beta + (\overline{\upsilon}_{ct} - \overline{\upsilon}_c). \tag{1.5}$$

Equations 1.3,1.4, and 1.5 together provide the basis for estimating $\tilde{\beta}_W$ by performing OLS on equation 1.5, which is known as the fixed effects estimator or the within estimator.[6]

In order to apply the fixed effect estimation to a dynamic model with pseudo-panel

[6] Adapted from STATA manual, section "xtreg, fe"

data, the main focus of this chapter, the conditions are slightly different from the actual panel data.[7]

$$\overline{y}_{c_{(t)}t} = \alpha_c + \gamma_c \overline{y}_{c_{(t-1)}t-1} + \overline{x}'_{c_{(t)}t}\beta_c + \overline{z}'_{c_{(t-1)}t-1}\delta_c + u_{c_{(t)}t}, \tag{1.6}$$

where $\bar{y}_{c_{(t)}t} = \sum \frac{y_{i,t}}{n_{c_{(t)}t}} \forall i \epsilon c$, $\bar{x}_{c_{(t)}t} = \sum \frac{x_{i,t}}{n_{c_{(t)}t}} \forall i \epsilon c$, $\bar{z}_{c_{(t-1)}t-1} = \sum \frac{x_{i,t-1}}{n_{c_{(t-1)}t-1}} \forall i \epsilon c_{(t-1)}$, $u_{c_{(t)}t} = \mu_{c_{(t)}} + \upsilon_{c_{(t)}t}$ and α_c is the cohort c specific intercepts when $c = 1, 2, \ldots, C$

The main difference is related to the lag variables. The value of any lag variable in the Model for cohort $c_{(t)}$ ($\bar{z}_{c_{(t)}t-1}$ or $\bar{y}_{c_{(t)}t-1}$) is unobservable. However, the value of that variable for the same cohort with random observations in the previous year ($\bar{z}_{c_{(t-1)}t-1}$ or $\bar{y}_{c_{(t-1)}t-1}$) is available. This situation creates unbiased estimators of cohort mean at time $t-1$ as the lag variable is estimated from a randomly distributed cohort from the previous period (McKenzie, 2004). In equation 1.6, $\overline{y}_{c_{(t)}t}$ is estimated average ratio of students in each cohort c who use marijuana at time t.

$\overline{x}'_{c_{(t)}t}$ is the matrix of estimated average ratios of independent variables in the Model at the time t. That consists of :

- The risk perception was reported,

- Having peers who use marijuana,

- Real family income and

- Living in metropolitan cities.

Also controls variables of age, gender, and risk attitude (See Table 1.16).

$\overline{z}'_{c_{(t-1)}t-1}$ the matrix of lagged variables consists of:

- The categorized average ratio of drug education in the past year,

- The average ratio of having a caring family in the past year,[8]

- The average ratio of students in each cohort who drank alcohol and smoked cigarettes during the past year, and

- The average ratio of students in each cohort who used marijuana in the past year

[7]Using pseudo-panel, implementing Hausman's test, and comparing the results of the fixed effect and random effect, lead the analysis to use fixed effect linear regression (See Table 1.16 and Table B.8)

[8]More information regarding caring family is in Table 1.12

$\bar{y}_{c_{(t-1)}t-1}$.

The error component $u_{c_{(t)}t}$ (equation 1.6) of a finite panel sample is correlated with $\bar{y}_{c_{(t-1)}t-1}$. As a result, least-squares estimations are biased (Inoue, 2008) while the fixed-effects estimator is inconsistent (Nickell, 1981). However, the fixed-effects estimator is asymptotically consistent when $n_{c(t)} \to \infty$ for fixed T (McKenzie, 2004). In contrast with the genuine panel case, $\overline{y}_{c_{(t+s)},t+s}$ for $S \geqslant 1$ is also uncorrelated with $\overline{y}_{c_{(t)}t-1} - \overline{y}_{c_{(t-1)}t-1}$ since different individuals are sampled each period, along with the assumption of cross-sectional independence (Inoue, 2008).

For every cohort c, assuming $n_{c_{(t)}} \to \infty$ for a fixed time, I follow Deaton (1985) and Moffitt (1993). Also, Verbeek (1995) took the same approach to have a consistent linear fixed-effects pseudo-panel for fixed C and T.

In this section, the asymptotic theory continues to hold, using the NSDUH database to create a pseudo-panel. First, the NSDUH has large cross-sectional observations (N) and small time (T) dimensions. Furthermore, the assumption that no one has ever used marijuana at age of 12 ($y_{i_{(t)}0} = 0$) helps avoid the initial condition problem.

$$\overline{y}_{c_{(t)}} = \alpha_c + \gamma_c \overline{y}_{c_{(t-1)}} + \overline{x}'_{c_{(t)}}\beta_c + \overline{z}'_{c_{(t-1)}}\delta_c + \overline{\mu}_c + \overline{v}_{c_{(t)}} \tag{1.7}$$

Subtracting equation 1.7 from equation 1.6 and writing it the matrix form,

$$\overline{y}_{c_{(t)}t} - \overline{y}_{c_{(t)}} = \gamma_c(\overline{y}_{c_{(t-1)}t-1} - \overline{y}_{c_{(t-1)}}) + (\overline{x}'_{c_{(t)}t} - \overline{x}'_{c_{(t)}})\beta_c + (\overline{z}'_{c_{(t-1)}t-1} - \overline{z}'_{c_{(t-1)}})\delta_c + (\overline{v}_{c_{(t)}t} - \overline{v}_{c_{(t)}}), \tag{1.8}$$

so

$$\begin{aligned} \tilde{y}_{c_{(t)}t} &= \gamma_c \tilde{y}_{c_{(t-1)}t-1} + \tilde{x}'_{c_{(t)}t}\beta_c + \tilde{z}'_{c_{(t-1)}t-1}\delta_c + \tilde{v}_{c_{(t)}t} \\ \tilde{y}_c &= \tilde{\mathit{\Gamma}}_c\theta_c + \tilde{v}_c, \qquad c = 1, 2, \ldots, C. \end{aligned} \tag{1.9}$$

$\theta_c = (\gamma_c, \beta'_c, \delta'_c)'$ indicates $(p+2) \times 1$ vector of all parameters for cohort c,

$\tilde{y}_c$, $\tilde{y}_{c,-1}$ and $\tilde{\varepsilon}_c$ are the vectors with elements $\overline{y}_{c_{(t)}t}$,

$\overline{y}_{c_{(t-1)}t-1}$, $\varepsilon_{c_{(t)}t}$, $\tilde{X}_c$ as matrix of $(t-1) \times p$ with rows $\overline{x}'_{c_{(t)}t}$ and $\overline{z}'_{c_{(t)}t}$ (p is number of independent variables.).

$\tilde{\Gamma}_c = (1, \tilde{y}_{c,-1}, \tilde{x}_c)$

Fixed effect is the OLS estimation of θ_c,

$$\hat{\theta}_c^{\,OLS} = (\tilde{\Gamma}'_c \tilde{\Gamma}_c)^{-1} \tilde{\Gamma}'_c \tilde{y}_c. \tag{1.10}$$

1.4.2 Principal Component Factor Analysis Estimation

A mixture of Principal Component Analysis (PCA) and Factor Analysis (FA) has been used in this section to produce "Caring family" information.

A PCA is a compression method used for dimension reduction involving technical mathematics procedures. As a result of PCA, a larger number of potentially correlated variables are reduced to a smaller number of uncorrelated variables while retaining most of their original information (Costello and Osborne, 2005; Torres-Reyna, 2009; Jolliffe, 2011). It is created by extracting the maximum variance from the observed variables and creating a linear combination of variables (Sweet and Grace-Martin, 2008).

In addition, factor analysis (FA) has a fundamental feature that rotates the axis to accommodate the actual data in the variable space of a multidimensional system (Sweet and Grace-Martin, 2008). The goal of factor rotation is to make the factors easier to interpret. However, when there is one factor for the analysis, rotation is not needed as the variables are not in a multidimensional space. Table 1.11 and Table 1.12 show the findings of the principal component factor analysis used to develop the "caring family" as a factor that carries information of four variables below for both the NSDUH cross-sections and pseudo-panel, respectively.

As "Caring family" is the only factor available for "Parents told you that you have done a good job during this year", "Parents checked if you have done your homework during this year", "Parents helped you with your homework during this year", and "Parents let you know they are proud of you during this year", rotation is not needed. However, to validate if the "caring family" is a good fit to be substituted in the Model instead of those four separate variables, the tests of Kaiser-Meyer-Olkin (KMO), Bartlett's and Cronbach's Alpha are suggested.

The Kaiser-Meyer-Olkin (KMO) (Kaiser and Rice, 1974) is the overall control of the Measuring of Sampling Adequacy (MSA) for a set of variables. KMO is a measure of correlation which clarifies if analyzing the correlation matrix has the potential for factor analysis (Arifin, 2017).

The value of the KMO test can be calculated as:[9]

$$KMO = \frac{trace(R^2) - p}{trace(R^2) + trace(Q^2) - 2p} \tag{1.11}$$

where the inverse of the observed correlation matrix is (R^{-1}), the anti-image correlation matrix is $Q = [(diagR^{-1})^{-1/2}]R^{-1}[(diagR^{-1})^{-1/2}]$ and p is number of variables. The quantity of calculated KMO in equation 1.11 is always in the closed interval [0,1] (Kaiser and Rice, 1974).

To analyze the KMO test, Kaiser and Rice (1974) suggest some criterion for KMO in Table 1.10 to compare with the calculated KMO from equation 1.11. Eventually, if KMO is below 0.5, using factor analysis is not recommended.

Table 1.10: KMO guideline (Kaiser and Rice, 1974)

Value	Interpretation
< 0.5	Unacceptable
0.5-0.59	Miserable
0.6-0.69	Mediocre
0.7-0.79	Middling
0.8-0.89	Meritorious
0.9-1.00	Marvelous

Bartlett's test of Sphericity was introduced by Bartlett (1951) to compare the observed correlation matrix with the identity matrix. This test is the first step in PCFA

[9]In STATA, it can be calculated by the post estimation command, estat kmo.

to recognize if the variables have enough correlation to be summarized in a few factor variables or not. In the null hypothesis, the variables are orthogonal and uncorrelated. The alternative hypothesis is that variables are sufficiently related, such that the correlation matrix deviates significantly from the identity matrix (Bartlett, 1951; Gorsuch, 1988).[10]

The formula for the Chi-square value of Bartlett's test of Sphericity is

$$\chi^2_{(\frac{P^2-P}{2})} = -((n-1) - \frac{2p-5}{6}) \times \log(det(R)) \tag{1.12}$$

where n is the number of observations, p is the number of variables and R is the correlation matrix and $\frac{P^2-P}{2}$ is the degrees of freedom.

The Cronbach's Alpha test is for estimating if the variables are internally consistent and reliable.[11] The threshold of 0.7 for Cronbach's Alpha is generally considered acceptable. However, similar to other statistical tools, this test also has some limitations. Cronbach's Alpha tests tend to have poorer results for categorical variables with a low score range (i.e., 0 - 1), and a threshold of 0.5 is generally accepted. Cronbach's Alpha formula is as follow:

$$\rho_T = \frac{k^2 \overline{\sigma}_{ij}}{\sigma_x^2}, \tag{1.13}$$

where

$\sigma_x^2 = \sum_{i=1}^k \sum_{j=1}^k \sigma_{ij} = \sum_{i=1}^k \sigma_i^2 + \sum_{i=1}^k \sum_{j\neq i}^k \sigma_{ij},$

$\overline{\sigma}_{ij} = \frac{\sum_{i=1}^k \sum_{j\neq i}^k \sigma_{ij}}{k(k-1)}$,

and k is the number of variables.

Following the analysis of each group of variables based on the introduced criteria and techniques, Table 1.11 and Table 1.12 present the results of the best-combination factor variable.

[10] The null hypothesis will be rejected if corresponding the p-value of Bartlett's test of Sphericity (the Chi-square test statistic) is less than 0.05

[11] Cronbach's Alpha is also known as Tau-equivalent reliability ρ_T.

The first section of Table 1.11 shows the information related to questions that are combined linearly to create a "caring family" factor variable. These categorical variables are scored from 0 to 1 in 4 categories of "Never=0", "Seldom=0.33", "Sometimes=0.67", and "Always=1." There is an acceptable Cronbach's Alpha of 0.75 and 0.79 for males and females, respectively. In addition, the KMO measure of sampling adequacy is very high for both genders.

Cumulative initial eigenvalues show that new factor variable (caring family) can reflect 57.3% and 61.7% of the information regarding those 4 variables for males and females, respectively. So this factor variable is an acceptable fit to be substituted to the Model instead of these four correlated variables. Finally, in the last section of Table 1.11, the linear functions that create caring family are mentioned for each gender.

Table 1.12 provides similar information regarding the NSDUH pseudo-panel. Both males and females have very high Cronbach's Alphas of 0.94 and 0.96, respectively. For both genders, the KMO is a good fit (0.79 for males and 0.8 for females). In addition, the cumulative initial eigenvalue is also high.

Considering the results of these tests in both Table 1.11 and Table 1.12, the "caring family" factor would make a suitable replacement in the Model for the four family characteristics discussed above.

Table 1.11: Principal component factor variable for NSDUH cross-sections(2002-2014)

Factor variable	Component variables	**Males Linear coefficient** (dy/dx)	**Females Linear coefficient** (dy/dx)
Caring Family	Parents let you know have done a good job in the past year	1.4248	1.2498
	Parents check if you've done your homework in the past year	0.9585	0.856
	Parents helped you with your homework in the past year	1.0068	0.941
	Parents let you know they are proud of you in the past year	1.378	1.2222
	Constant	−3.8085	−3.3034
	Cronbach's Alpha (Reliability statistics test)	0.75	0.79
	Kaiser-Meyer-Olkin Measure of Sampling Adequacy	0.66	0.69
	Barrtlett's and Test of Sphericity (significant level)	0.000	0.000
	Cumulative initial Eigenvalues	57.3%	61.7%

Refer to the linear combination of the variables (dy/dx) and their value between zero and one, calculated formulas are:

Males: $caringfamily = (-3.8085) + (1.4248)(v.1) + (0.9585)(v.2) + (1.0068)(v.3) + (1.378)(v.4)$
And the value of caring Family for males is between -3.8085 and 0.9596

Females: $caringfamily = (-3.3034) + (1.2498)(v.1) + (0.856)(v.2) + (0.941)(v.3) + (1.2222)(v.4)$
And the value of caring Family for females is between -3.3034 and 0.9656

Table 1.12: Principal component factor variable for NSDUH pseudo-panel(2002-2014)

Factor variable	Component variables	**Males Linear coefficient** (dy/dx)	**Females Linear coefficient** (dy/dx)
Caring Family	(v.1) Parents let you know have done a good job in the past year	5.4652	3.8768
	(v.2) Parents check if you've done your homework in the past year	3.5808	2.6821
	(v.3) Parents helped you with your homework in the past year	3.1334	2.4833
	(v.4) Parents let you know they are proud of you in the past year	5.5198	3.9218
	Constant	−14.0963	−9.8016
	Cronbach's Alpha (Reliability statistics test on standardized items)	0.94	0.96
	Kaiser-Meyer-Olkin Measure of Sampling Adequacy	0.79	0.8
	Barrtlett's and Test of Sphericity (significant level)	0.000	0.000
	Cumulative initial Eigenvalues	84.3%	88.3%

Refer to the linear combination of the variables (dy/dx) and their value between zero and one, calculated formulas are:

Males: $caringfamily = (-14.0963) + (5.4652)(v.1) + (3.5808)(v.2) + (3.1334)(v.3) + (5.5198)(v.4)$
And the value of caring Family for males can be between −14.0963 and 3.0603.
However the rang in the sample is between −2.9306 and 1.9631

Females: $caringfamily = (-9.8016) + (3.8768)(v.1) + (2.6821)(v.2) + (2.4833)(v.3) + (3.9218)(v.4)$
And the value of caring Family for females can be between −9.8016 and 3.1624
However the rang in the sample is between −2.919 and 2.1079

1.5 Empirical Results

This chapter examines the effect of school-based drug education on the use of marijuana with pseudo-panel data containing 900 cohorts and an average of 152 individuals per cohort every year, which is constructed from 136,555 respondents in the NSDUH cross-sectional annual surveys. The large numbers of observations per cohort allow this study to ignore measurement errors similar to Moffitt (1993), Verbeek (1995), McKenzie (2004), and Bernard et al. (2011) and estimates a fixed effect linear regression for the pseudo-panel similar to an actual panel.

In general, other studies investigated the effect of specific school-based drug prevention programs at specific schools concerning drug consumption (Onrust et al., 2016; Fletcher et al., 2008; Birkeland et al., 2005; Bond et al., 2004; Wilson et al., 2001; Ennet et al., 1994; Botvin et al., 1990a). However, the current study's novelty is analyzing reported drug education by students aged 12 to 17 from a national US survey NSDUH by creating a pseudo-panel and the ability to consider dynamics analysis.

Respecting the final Models in Table 1.16 and Table 1.17, the hierarchical regression technique helps to choose the best-fitted Model with the lowest AIC (Akaike's Information Criterion) (Pan, 2001) and BIC (the Bayesian Information Criterion)(Dziak et al., 2012) and the highest R-squared.

AIC and BIC are generally consistent and lead to the same conclusion. However, in this study, where AIC and BIC lead to different results, AIC takes precedence.[12] AIC and BIC are calculated as follows:

$$
\begin{aligned}
AIC &= -2\ln(l) + 2K, \\
BIC &= -2\ln(l) + ln(N)\,k,
\end{aligned}
\tag{1.14}
$$

where

$\ln(l)$ is the log-likelihood of the Model,

k is the degrees of freedom (or independently adjusted parameters) in the Model, and

N is the number of observations (total sample size).

As a result of the Hausman test and Guillerm's approach in 2017 concerning pseudo-panels, both main Models in Table 1.16 and Table 1.17 are estimated based on fixed-effect linear regressions. Despite those reasons, to choose the best Model, many Models are estimated, of which the most important ones are listed in Appendix B. Table B.6 compares fixed and random-effect estimates, and Table B.9 compares pools versus fixed-effect estimates.

Table 1.16 illustrates how six Models incorporate independent variables step-by-step into a fixed-effect linear Model to estimate the direct effect of school-based drug education on marijuana use. Beside, Table 1.17 shows the results of four hierarchical regression estimations to analyze the effect of school-based drug education on the reported risk perception by adolescents.

[12]The best Model is the one that minimizes both AIC and BIC. In summary, BIC shows the danger of underfitting, AIC represents the danger of overfitting, and the penalty for adding parameters is higher in BIC than in AIC (Brewer et al., 2016; Chatterjee and Hadi, 2012; Dziak et al., 2012; Kuha, 2004).

1.5.1 School-Based Drug Prevention Programs

The NSDUH includes a dummy variable for drug education in schools, which is transformed into a continuous variable during the pseudo-panel creation. The school-based drug education in the pseudo-panel is a variable between 0 and 1 that indicates the percentage of individuals in each cohort who received school-based drug education.

Figure 1.2: The percentage of individuals per cohort in each age group (12 to 17) who received school-based drug education in the NSDUH pseudo-panel

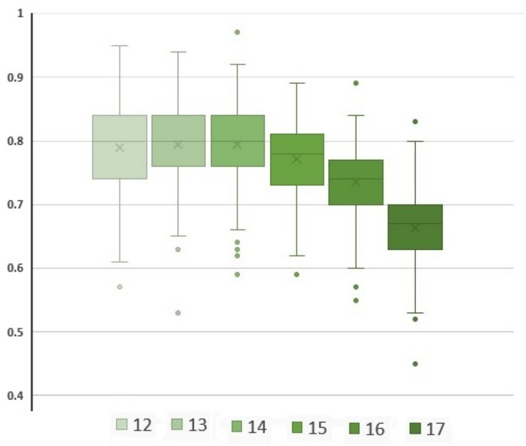


Figure 1.2 shows the maximum, minimum, and mode of the percentage of students per cohort who have received drug education at school in each age group. The proportion of individuals receiving school-based drug education is highest in cohorts of 13- and 14-year-olds, while the rate decreases as respondents age.

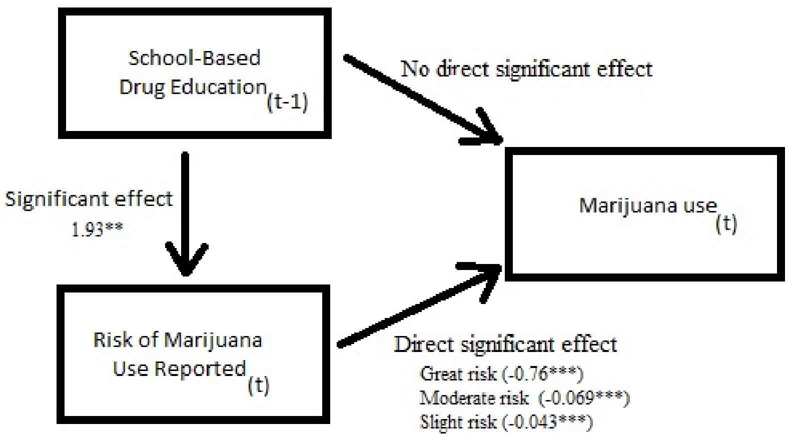
Table 1.16 shows estimated Models that analyze the effect of school-based drug education on marijuana use. The estimated Models in Table 1.17 describe the influences of similar factors in Table 1.16 on reported risk perception. Table 1.16 shows that school-based drug education does not significantly affect marijuana use; however, reported risk perception substantially affects marijuana use, While school-based drug education significantly impacts the reported risk perception in Table 1.17.

In Table 1.18, Model 1 analyzes the effect of school-based drug education on marijuana use without reporting risk perception, and the result shows that school-based drug education will decrease marijuana use. Model 2 illustrates the effect of school-based drug education on reported risk perception. In this case, the effect is positive and highly significant.

Model 3 is similar to Model 1 with the addition of reported risk perceptions. The effects of school-based drug education are reduced by approximately 50% when using this Model, and all levels of reported risk perception are associated with significant reductions in marijuana use. Adding marijuana, alcohol, and cigarette habits to Model 4 decreases the effect of reported risk perception at all levels; however, it still significantly affects marijuana use.

According to Table 1.18, school-based drug education reduces marijuana use indirectly by increasing reported risk perceptions (see Figure 1.3).

Figure 1.3: The direct and indirect effects of school-based drug education on marijuana use



Considering Table 1.16, adolescents' reported risk perception has a significant negative impact on marijuana use. Returning to the results in Model 6, there are 4 percentage points fewer marijuana users in the cohort that believes marijuana use is slightly risky compared to those who think marijuana use is not risky. Furthermore, there is a 6.9 percentage points decrease in the percentage of users in groups who believe it has a moderate risk and a 7.6 percentage points decrease in those who think it has a great risk. Consequently, despite the fact that school-based drug education has no significant effect directly on marijuana use, it increases marijuana use risk by roughly two degrees on average among adolescents. Therefore, school-based drug education decreases marijuana use based on the reported risk perception. For example, drug education at school can change a cohort's perception from no-risk to moderate-risk, which results in a decrease of around 6.9 percentage points in marijuana users, or from slight risk to great risk, which results in a 3.3 percentage points reduction in marijuana users.[13]

1.5.1.1 How Long Does the Effect of School-Based Drug Prevention Programs Remain?

One of the most important questions, which is also discussed in previous studies (Soole et al., 2008; Rosenbaum, 2007; Bond et al., 2004; White and Pitts, 1998), is how long the school-based drug education remains effective. This study estimates the final Models of Table 1.16 and Table 1.17 with three different timelines demonstrated in Table B.11 and Table B.12 to answer this question. The first Models in both tables analyze the effect of last year's school-based drug education. In the second Model, the impact of school-based drug education from two years ago is added to the estimation, and in the third Model, school-based drug education from three years ago is also included.

A one-year lag is found to be the most appropriate Model based on the results from Table B.11 and Table B.12 (see Table 1.13). It is in accordance with Čurová et al. (2021), Orosová et al. (2020), Gabrhelik et al. (2012), and Vigna-Taglianti et al. (2009) regarding the short-term effectiveness of school-based drug prevention programs.

[13]The reported risk perception by adolescents is a categorical variable in four degrees of 0=No Risk, 1=Slight Risk, 2=Moderate Risk, and 3=Great Risk.

Table 1.13: Examining the effect of school-based drug education over time

	Final Model one-year lag	two years lags	three years lags
From Table B.11			
Effect of drug education on marijuana use			
Drug info$_{(t-1)}$=1	0.042 [0.059]	0.055 [0.102]	-0.041 [0.178]
Drug info(t-2)=1		0.01 [0.117]	-0.062 [0.222]
Drug info(t-3)=1			0.06 [0.190]
R-squared Overall	0.877	0.81	0.389
From Table B.12			
Effect of drug education on the reported risk perception			
Drug info$_{(t-1)}$=1	1.930** [0.687]	2.841* [1.195]	-0.865 [2.401]
Drug info(t-2)=1		1.784 [1.363]	2.19 [2.243]
Drug info(t-3)=1			1.775 [2.538]
R-squared Overall	0.395	0.266	0.068

- $^{*}p < 0.05,^{**}p < 0.01,^{***}p < 0.001$
- The standard deviation is inside the bracket [...]

1.5.1.2 Is the Time of Delivering School-Based Drug Education Important?

On average, aging among adolescents increases the percentage of individuals in each cohort who use marijuana by 1.3 percentage points regardless of school-based drug education (see Table 1.16, Model 6). On the other hand, aging reduces the reported risk perception significantly by 0.327 levels on average (see the first section of Table 1.14). Based on the results in second part of Table 1.14, 16-year-old cohorts who received school-based drug education, on average reported 0.37 lower levels of the perceived risk of marijuana use than the other cohorts. In conclusion, the time for delivering school-based drug education is important, and for the highest results, drug education should start before age 15.

Table 1.14: Age and school-based drug education interaction effect on the reported risk perception

	Final Model
From Table 1.17	
Drug info$_{(t-1)}$=1	1.930**
	[0.687]
Age	-0.327^{***}
	[0.048]
Age# Drug info$_{(t-1)}$=1	-0.109^{**}
	[0.041]
From Table B.16	
Drug info$_{(t-1)}$=1	0.524*
	[0.204]
Age	
13	*Reference*
14	-0.599^{**}
	[0.201]
15	-0.941^{***}
	[0.230]
16	-1.075^{***}
	[0.203]
17	-1.469^{***}
	[0.201]
Drug info$_{(t-1)}$=1#13	*Reference*
Drug info$_{(t-1)}$=1#14	0.153
	[0.201]
Drug info$_{(t-1)}$=1#15	-0.177
	[0.216]
Drug info$_{(t-1)}$=1#16	-0.370^{*}
	[0.175]
Drug info$_{(t-1)}$=1#17	-0.33

- $^{*}p < 0.05$,$^{**}p < 0.01$,$^{***}p < 0.001$
- The standard deviation is inside the bracket [...]

Using the estimated models in Table B.6 and Table B.7, Table 1.15 summarizes the effect of reported marijuana risk in different age categories on marijuana use. Compared to other cohorts, 15- and 16-year-old students use more marijuana at all levels of risk perception.

Table 1.15: Age and the reported risk of marijuana interaction effect on marijuana use

	Final Model
From Table B.6 and Table B.7	
The reported risk perception	
No risk	Reference
Slight risk	-0.194^{***}
	[0.071]
Moderate risk	-0.209^{**}
	[0.08]
Great risk	-0.255^{**}
	[0.081]
The reported risk perception#Age	
Slight risk#15	0.063^{**}
	[0.024]
Moderate risk#15	0.079^{***}
	[0.020]
Moderate risk#16	0.043^{*}
	[0.018]
Great risk#15	0.054^{***}
	[0.015]
Great risk#16	0.037^{***}
	[0.013]

$^{*}p < 0.05, ^{**}p < 0.01, ^{***}p < 0.001$
The standard deviation is inside the bracket [...]

Figure 1.4 shows that with aging, the perception of the risk of using marijuana decreases.

In Figure 1.5, the percentage of marijuana users among cohorts increases with aging.

1.5.1.3 How Does School-Based Drug Education Effect Differ by Demographics, Peers, and Families?

According to previous discussions, school-based drug education only influences marijuana use through the channel of the reported risk perception among adolescents. Based on these results, school-based drug education's impact on the reported risk perception is not affected by gender, caring family, real family income, living in metropolitan areas, or having peers who use marijuana. However, Black and Hispanic cohorts who received school-based drug education reported a lower risk perception. On average,

Figure 1.4: The reported risk perception among cohorts 12 to 17
NSDUH pseudo-panel (2002-2014)

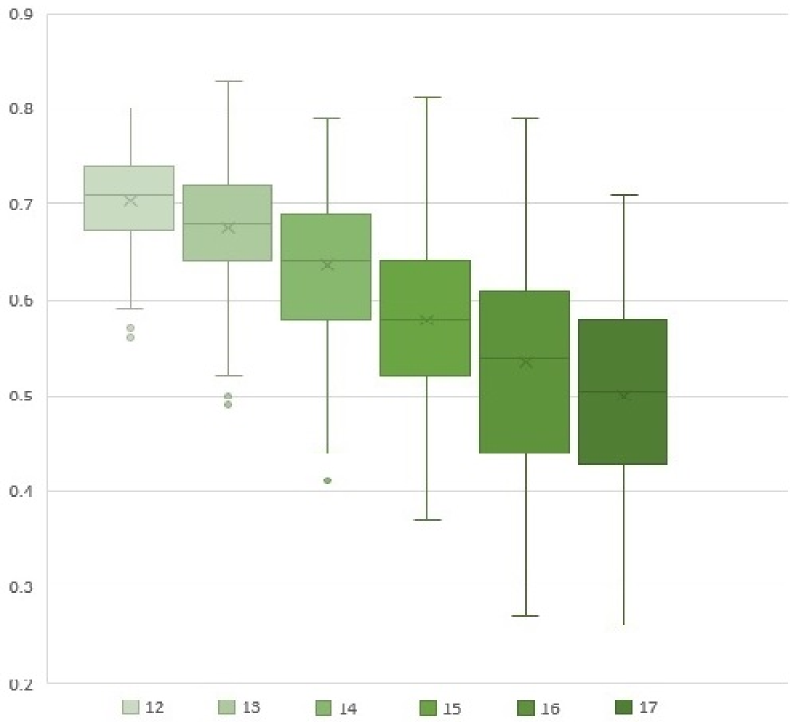
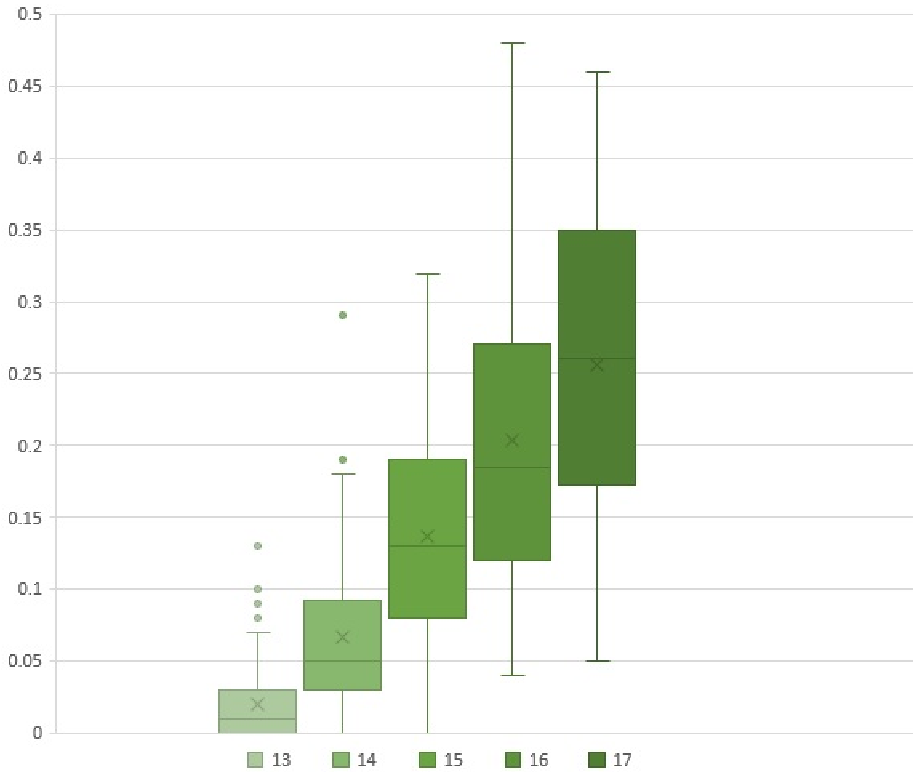


Figure 1.5: Percentage of marijuana users per cohort at ages 13 to 17
NSDUH pseudo-panel (2002-2014)



school-based drug education is less effective in reducing marijuana use among Black and Hispanic students than among non-Black, non-Hispanic students (See Table 1.17 and Table B.17).

Table 1.16: Marijuana use and drug education (2002-2014)
fixed effect linear weighted NSDUH pseudo-panel data

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Drug info$_{(t-1)}$=1	-0.055	-0.019	-0.004	-0.059	0.044	0.042
	[0.088]	[0.073]	[0.072]	[0.065]	[0.059]	[0.059]
Age	0.056^{***}	0.043^{***}	0.034^{***}	0.029^{***}	0.013^{***}	0.013^{***}
	[0.005]	[0.004]	[0.004]	[0.004]	[0.004]	[0.004]
Female # Drug info$_{(t-1)}$=1	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*
Male # Drug info$_{(t-1)}$=1	-0.011	-0.01	-0.015	-0.015	-0.013	-0.013
	[0.014]	[0.012]	[0.012]	[0.010]	[0.010]	[0.010]
Non-Black, non-Hispanic # Drug info$_{(t-1)}$=1	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*
Black# Drug info$_{(t-1)}$=1	-0.039^{**}	-0.030^{*}	-0.028^{*}	-0.015	-0.016	-0.016
	[0.015]	[0.015]	[0.014]	[0.013]	[0.013]	[0.013]
Hispanic# Drug info$_{(t-1)}$=1	-0.034^{*}	-0.032^{*}	-0.031^{*}	-0.016	-0.016	-0.015
	[0.016]	[0.016]	[0.014]	[0.012]	[0.013]	[0.013]
Age# Drug info$_{(t-1)}$=1	0.007	0.004	0.003	0.006	-0.001	-0.001
	[0.005]	[0.004]	[0.004]	[0.004]	[0.003]	[0.003]
risk perception reported						
No risk	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*
Slight risk		-0.069^{***}	-0.066^{***}	-0.058^{***}	-0.043^{***}	-0.043^{***}
		[0.006]	[0.006]	[0.005]	[0.005]	[0.005]
Moderate risk		-0.104^{***}	-0.100^{***}	-0.093^{***}	-0.070^{***}	-0.069^{***}
		[0.009]	[0.009]	[0.007]	[0.006]	[0.006]
Great risk		-0.101^{***}	-0.097^{***}	-0.100^{***}	-0.076^{***}	-0.076^{***}
		[0.012]	[0.012]	[0.009]	[0.007]	[0.007]
Caring family$_{(t-1)}$			-0.083^{***}	-0.068^{***}	-0.022	-0.022
			[0.017]	[0.016]	[0.012]	[0.012]
Peers use marijuana						
None of them				*Reference*	*Reference*	*Reference*
A few of them				-0.013^{**}	0.007^{*}	0.007^{*}
				[0.004]	[0.003]	[0.003]
Most of them				0.025^{**}	0.039^{***}	0.038^{***}
				[0.009]	[0.006]	[0.007]
All of them				0.093^{***}	0.092^{***}	0.092^{***}
				[0.023]	[0.022]	[0.022]
Marijuana use$_{(t-1)}$					0.184^{***}	0.184^{***}
					[0.048]	[0.047]
Drink alcohol$_{(t-1)}$					0.213^{***}	0.212^{***}
					[0.030]	[0.029]
Smoke cigarette$_{(t-1)}$					0.002	0.002
					[0.055]	[0.055]
Real family income						0.002
						[0.005]
Live in metropolitan city=1						0
						[0.003]
Constant	-0.733^{***}	-0.462^{***}	-0.324^{***}	-0.252^{***}	-0.1	-0.111
	[0.079]	[0.074]	[0.074]	[0.069]	[0.059]	[0.065]
N	900	900	900	900	900	900
AIC	-3400	-3700	-3700	-3900	-4100	-4100
BIC	-3400	-3700	-3700	-3800	-4000	-4000
sigma_u	0.058	0.046	0.035	0.031	0.022	0.023
sigma_e	0.042	0.035	0.035	0.032	0.028	0.028
rho	0.657	0.632	0.509	0.481	0.379	0.388
R-squared						
Overall	0.557	0.705	0.794	0.831	0.879	0.877

- $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$
- The standard deviation is inside the bracket [...]

Table 1.17: Risk of marijuana use perception and drug education (2002-2014) fixed effect linear weighted NSDUH pseudo-panel Data

	Model 1	Model 2	Model 3	Model 4
Drug info$_{(t-1)}$=1	2.058^{**}	1.958^{**}	1.909^{**}	1.930^{**}
	[0.682]	[0.692]	[0.672]	[0.687]
Age	-0.410^{***}	-0.353^{***}	-0.327^{***}	-0.327^{***}
	[0.034]	[0.044]	[0.048]	[0.048]
Female # Drug info$_{(t-1)}$=1	*Reference*	*Reference*	*Reference*	*Reference*
Male # Drug info$_{(t-1)}$=1	-0.085	-0.059	-0.045	-0.048
	[0.116]	[0.115]	[0.116]	[0.116]
Non-Black, non-Hispanic # Drug info$_{(t-1)}$=1	*Reference*	*Reference*	*Reference*	*Reference*
Black# Drug info$_{(t-1)}$=1	-0.362^{*}	-0.374^{*}	-0.333^{*}	-0.336^{*}
	[0.150]	[0.155]	[0.164]	[0.164]
Hispanic#Drug info$_{(t-1)}$=1	-0.399^{*}	-0.403^{*}	-0.387^{*}	-0.389^{*}
	[0.176]	[0.177]	[0.185]	[0.184]
Age# Drug info$_{(t-1)}$=1	-0.116^{**}	-0.109^{**}	-0.108^{**}	-0.109^{**}
	[0.041]	[0.042]	[0.040]	[0.041]
Caring family$_{(t-1)}$		0.492^{*}	0.546^{*}	0.549^{*}
		[0.225]	[0.219]	[0.220]
Peers use marijuana				
None of them			*reference*	*reference*
A few of them			-0.12	-0.116
			[0.063]	[0.062]
Most of them			-0.076	-0.071
			[0.115]	[0.113]
All of them			-0.447^{*}	-0.442^{*}
			[0.185]	[0.185]
Real family income				-0.012
				[0.081]
Live in metropolitan city=1				-0.028
				[0.057]
Constant	7.521^{***}	6.645^{***}	6.351^{***}	6.419^{***}
	[0.573]	[0.733]	[0.769]	[0.852]
N	900	900	900	900
AIC	946.978	941.701	936.89	940.311
BIC	975.792	975.318	984.914	997.94
sigma_u	0.719	0.689	0.686	0.685
sigma_e	0.469	0.467	0.465	0.466
rho	0.702	0.685	0.685	0.684
R-squared				
Overall	0.346	0.387	0.394	0.395

- $^{*}p < 0.05,^{**} p < 0.01,^{***} p < 0.001$
- The standard deviation is inside the bracket [...]

Table 1.18: Marijuana use, drug education and risk perception (2002-2014) fixed effect linear weighted NSDUH pseudo-panel data

	Model 1	Model 2	Model 3	Model 4
	Marijuana use	Risk perception	Marijuana use	Marijuana use
Drug info$_{(t-1)}$=1	-0.126	1.930^{**}	-0.062	0.042
	[0.071]	[0.687]	[0.065]	[0.059]
Age	0.039^{***}	-0.327^{***}	0.030^{***}	0.013^{***}
	[0.004]	[0.048]	[0.004]	[0.004]
Female# Drug info$_{(t-1)}$=1	*Reference*	*Reference*	*Reference*	*Reference*
Male# Drug info$_{(t-1)}$=1	-0.015	-0.048	-0.015	-0.013
	[0.011]	[0.116]	[0.010]	[0.010]
Non-Black, non-Hispanic# Drug info$_{(t-1)}$=1	*Reference*	*Reference*	*Reference*	*Reference*
Black# Drug info$_{(t-1)}$=1	-0.012	-0.336^{*}	-0.015	-0.016
	[0.013]	[0.164]	[0.013]	[0.013]
Hispanic# Drug info$_{(t-1)}$=1	-0.006	-0.389^{*}	-0.015	-0.015
	[0.012]	[0.184]	[0.012]	[0.013]
Age# Drug info$_{(t-1)}$=1	0.010^{*}	-0.109^{**}	0.006	-0.001
	[0.004]	[0.041]	[0.004]	[0.003]
Caring family$_{(t-1)}$	-0.090^{***}	0.549^{*}	-0.066^{***}	-0.022
	[0.018]	[0.220]	[0.016]	[0.012]
Peers use marijuana				
Non of them	*Reference*	*Reference*	*Reference*	*Reference*
A few of them	-0.018^{***}	-0.116	-0.013^{**}	0.007^{*}
	[0.004]	[0.062]	[0.004]	[0.003]
Most of them	0.026^{**}	-0.071	0.025^{**}	0.038^{***}
	[0.009]	[0.113]	[0.008]	[0.007]
All of them	0.116^{***}	-0.442^{*}	0.094^{***}	0.092^{***}
	[0.023]	[0.185]	[0.022]	[0.022]
Real family income	0.014^{*}	-0.012	0.005	0.002
	[0.007]	[0.081]	[0.006]	[0.005]
Live in metropolitan city=1	-0.001	-0.028	-0.003	0
	[0.004]	[0.057]	[0.004]	[0.003]
risk perception reported				
No risk			*Reference*	*Reference*
Slight risk			-0.058^{***}	-0.043^{***}
			[0.005]	[0.005]
Moderate risk			-0.093^{***}	-0.069^{***}
			[0.007]	[0.006]
Great risk			-0.100^{***}	-0.076^{***}
			[0.009]	[0.007]
Marijuana use$_{(t-1)}$				0.184^{***}
				[0.047]
Drink alcohol$_{(t-1)}$				0.212^{***}
				[0.029]
Smoke cigarette$_{(t-1)}$				0.002
				[0.055]
Constant	-0.525^{***}	6.419^{***}	-0.277^{***}	-0.111
	[0.080]	[0.852]	[0.075]	[0.065]
N	900	900	900	900
AIC	-3600	940.3	-3900	-4100
BIC	-3600	997.9	-3800	-4000
sigma_u	0.039	0.685	0.031	0.023
sigma_e	0.037	0.466	0.032	0.028
rho	0.52	0.684	0.488	0.388
R-squared				
Overall	0.753	0.395	0.828	0.877

- $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$
- The standard deviation is inside the bracket [...]

1.6 Conclusions

Drug abuse is an important topic in labour economics, public health, and social security. In addition, smoking, drinking, violence, dangerous sexual practices, and delinquency are strongly correlated with it among adolescents at an exceptional level (Flay and Collins, 2005; Jessor and Jessor, 1977). Therefore, randomized prevention programs have been routinely conducted at schools since the early 1980s to reduce these high-risk behaviours.

Following these program trials, a series of studies are conducted to assess the effectiveness of prevention programs in diminishing risky behaviours. Shackleton et al. (2016) and Perry et al. (2003), conclude that school-based drug prevention programs do not decrease the drug use. However, Bond et al. (2004), Wilson et al. (2001), Botvin et al. (1995), Gerstein and Green (1993), Hansen (1992), and Botvin et al. (1990b) find that drug prevention programs effectively reduce substance abuse at schools. Researchers such as Wilson et al. (2001) and Tobler et al. (1999) believe the lack of effectiveness of school-based drug prevention programs can be linked to inefficient examination methods and incomparable samples.

This study examines the impact of direct school-based drug education on marijuana use without considering any particular school or program. The sample was taken from the National Survey on Drug Use and Health (NSDUH) between 2002 and 2014, which covers the entire United States, not just a particular state or region. Also, students are asked if they received any drug education at school. As a result, direct drug education programs are analyzed since students are aware of receiving drug education but do not know the name of the prevention program.

An additional unique characteristic of this study is its use of a pseudo-panel methodology, which allows for the examination of students' attitudes toward drug education at schools over time, categorized by race, gender, and risk attitude.

According to weighted fixed-effect linear regression results in Table 1.18, school-based drug education does affect marijuana use through the reported marijuana risk perception. Thus, school-based drug education increases the students' reported marijuana risk perceptions and the reported marijuana risk perception decreases marijuana use.

The estimates in Table 1.16 provide a basis for developing the second weighted fixed-effect linear Models in Table 1.17 to analyze factors contributing to the reported marijuana risk perception among students. The estimated Model 4 in Table 1.17 shows a significant increase in the reported marijuana risk perception due to school-based drug education. The bottom line is that school-based drug education reduces marijuana use by increasing the reported marijuana risk perception.

In contrast to Perry et al. (2003), school-based education benefits both girls and boys equally. However, school-based drug education has a smaller effect on students' reported risk perception when they are 16 years old. Also, despite knowing that marijuana is dangerous, students between the ages of 15 and 16 appear to care less about the health risks associated with marijuana use and more about following their peers, which is consistent with Onrust et al. (2016)'s finding. Additionally, Black and Hispanic students who received school-based drug education reported a lower perceived risk of marijuana use than their non-Black and non-Hispanic peers.

Similar to previous studies on this topic, measurement error is a potential limitation that should be addressed, particularly regarding the creation of a pseudo-panel, the reported risk perception, and information about marijuana use itself. However, the problem was reduced by considering specific assumptions such as the initial conditions of not using marijuana before the age of 12 and carefully selecting the cohorts according to the requirements considered by Deaton (1985), Verbeek and Nijman (1992), and Guillerm (2017).

Chapter 2

Heterogeneous Effects of Marijuana Use on Educational Attainment

Abstract

We analyze transitions into the use of marijuana jointly with grade transitions using data from the 1997 cohort of the National Longitudinal Survey of Youth. We allow for correlated unobserved heterogeneity that impact both transitions within a discrete-time hazard framework. We estimate the impacts at different grade levels and find that they vary significantly. Average marginal effects indicate that using marijuana reduces next year's grade transition by 9.6 percentage points while in high school and by 2.3 percentage points while in college. The negative effects are larger for male youth and for students with weaker family background. They are also robust to different consumption levels.

JEL Code: I12,I21

Keywords: marijuana; education; grade transitions; discrete-time hazard; unobserved heterogeneity

2.1 Introduction

A large body of research has documented a negative association between early marijuana use and educational attainment (e.g., (Beverly et al., 2019; Melchior et al., 2017; Verweij et al., 2013; McCaffrey et al., 2010; Van Ours and Williams, 2009; Chatterji, 2006; Fergusson et al., 2003; Register et al., 2001; Bray et al., 2000; Yamada et al., 1996)). However, there is no general agreement about the causal effect of early marijuana use on educational outcomes. There may exist a causal relationship where marijuana use impacts brain functions, reduces learning and ultimately results in lower educational achievements. However, there is also a possibility or reversed causality where poor educational outcomes leads to marijuana use. A third possibility is that marijuana use and educational attainment are not directly related but depend on common factors, such as attitudes, family background, peers and time preferences.

Identifying the causal effects of marijuana use on educational outcomes is consequently challenging. First, there are numerous potential common factors, many of which are not available in observational data. Second, recent papers in the economics of education literature have documented the importance of selection on unobservables when modeling educational outcomes. Third, approaches using linear instrumental variables (IV) to address the endogeneity of marijuana use are less likely to be successful when outcomes are generated by sequential choices. Further, even if linear regression models were able to deal with this dynamic selection, obtaining valid instruments is difficult. Many studies using IV to estimate the effects of marijuana use on education rely on weak instruments and fail to reject the assumption of endogeneity (Chatterji, 2006; Register et al., 2001; Bray et al., 2000).

In this paper, we describe and estimate transitions into marijuana use for a recent cohort of American youth (the 1997 cohort of the National Longitudinal Survey of Youth, NLSY) and how they relate to observed and unobserved individual and family characteristics. We follow Van Ours and Williams (2009) and analyze how marijuana use impacts educational attainment using a bi-variate duration model. The time until marijuana initiation and educational investments are modeled jointly and unobserved heterogeneity affecting time until marijuana use is potentially correlated with unobserved heterogeneity that determines grade transitions. This is important as it addresses the endogeneity of marijuana use. Moreover, unobserved heterogeneity is assumed to follow a discrete distribution with four points of support.

Our paper addresses some important shortcomings in the previous literature. Unlike Van Ours and Williams (2009), we have detailed information on grade transitions in school. Moreover, respondents in our sample were asked about substance use at a much younger age (in some cases at age 12) than in their paper and we argue that the issue of recall errors is less serious in our paper. Also, unlike Beverly et al. (2019), Verweij et al. (2013),McCaffrey et al. (2010) and Van Ours and Williams (2009), we use data from a nationally representative longitudinal study that allow us to estimate the effects of early marijuana use on educational attainment from age 16 and onwards. Unlike any of the papers in the related literature, we estimate the impacts of marijuana use at different grade levels and find that they vary significantly. Another important contribution is the analysis of heterogeneity in effects across different individuals and the importance of unobserved heterogeneity.

Our results suggest that marijuana use has a significant, negative effect on grade transitions, both in high school and in college. The negative effect declines with highest grade completed and is largest in high school. We show that a model without controls for correlated unobserved heterogeneity severely exaggerates the negative effects from marijuana use at all grade levels. Similar to McCaffrey et al. (2010),we find that omitting consumption of tobacco and alcohol generates larger negative effects. The average marginal effects from our preferred model specification indicate that using marijuana reduces next year's grade transition by 9.6 percentage points while in high school and by 2.3 percentage points while in college. The corresponding effects in the naive, single spell model are 23.7 and 8.1 percentage points, respectively. We estimate larger negative effects for males (11.6 percentage points) than for females (7.4 percentage points) in high school and larger negative effects for Hispanic students than other students. We show that the negative effect of marijuana use is stronger for youth from weaker family backgrounds (low income, single mothers and teenage mothers) and that unobserved heterogeneity explains a substantial proportion of the variation in marijuana initiation and grade transitions, even after controlling for a rich set of observed characteristics. Further, we show that adding peer effects significantly contribute to marijuana initiation and educational attainment. However, unlike the finding in McCaffrey et al. (2010), incorporating peer effects in the set of covariates for grade transitions does not change the estimated, negative marginal effects of marijuana use on education. We also demonstrate how the effect of marijuana use on grade transitions vary across unobserved heterogeneity types. The effect is between -0.13 and -0.15 for just over half of the sample but much smaller for the rest, -0.028 and -0.056 for two types, each representing about 23 percent of the sample. Finally, we show that the results are robust to different consumption levels. The marginal effects of marijuana use on grade transitions in high school when we define consumption based on any use (at least once during the month preceding the interview) are similar to those obtained when we record consumption only if the person used it 10 times or more that same month.

The rest of the paper is organized as follows. In the next section, we provide a review of the existing literature on the effect of marijuana use on education. The data is described in Section 3 and the econometric model is presented in Section 4. The results are discussed in Section 5 and Section 6 concludes the paper with a brief summary.

2.2 Literature

As mentioned above, there has been considerable research devoted to the link between early marijuana use and educational attainment. Some of the papers are not in economics and ignore important dimensions and recent advances in the economics of education (such as endogeneity and selectivity). Most recognize the potential endogeneity of marijuana use but fail to convincingly address the issue. Below we review a selection of related studies.

Two early studies on this topic are Bray et al. (2000) and Yamada et al. (1996). In the former paper, the impact of four categories of substances (alcohol, cigarettes, marijuana, and other illicit drugs) on years of education is analyzed. Using data from four waves of a longitudinal survey of students in a southeastern U.S. public school system, they report that marijuana users are 2.3 times more likely to drop out of high school than non-users. A similar result was reported in Yamada et al. (1996) who, using data from the the 1979 cohort of the National Longitudinal Study of Youth (NLSY79), show that frequent marijuana use in grade 12 reduces the probability of graduating from high school by 5.6 percentage points.

Register et al. (2001) also use the NLSY79 to investigate the effects of early cannabis use on educational achievement. However, their outcome variable is years of schooling instead of high school graduation. Early cannabis use is represented by an indicator for any consumption of marijuana before the age of 18. They use survey responses to questions on religious activities and state marijuana laws as instruments. Their results suggest that using cannabis before age 18 reduces educational attainment with one year for white respondents but has no significant impact on years of schooling for Black or Hispanic students.

Chatterji (2006) use data from the 1988 National Education Longitudinal Study (NELS) to estimate the association between marijuana and cocaine use during high school and completed years of schooling. To address the endogeneity issue, the paper employs an IV approach using state-level substance use policies. The IV results indicate no significant effect of marijuana use, this is a result of inflated standard errors from using weak instruments, while the corresponding OLS estimates are negative and significant. Surprisingly, the negative effect from using marijuana in grade 12 is larger than the effect from using marijuana in grade 10.

A more relevant study for our paper is Van Ours and Williams (2009). They use a sample of 25-50 year old respondents to the 2001 Australian National Drug Strategy Household Survey to analyze the association between onset of marijuana use and age when leaving school. Since the data is cross-sectional, they use information from retrospective questions on how old respondents were when they first used marijuana. Unfortunately, they do not have detailed information on age when leaving school and instead infer that age based on the highest qualification the person has completed. They derive a bi-variate mixed proportional hazard model, where the transition rates into marijuana use and out of school are jointly estimated. Their results show that the probability that a female student leaves school is 51 percent higher for a cannabis user compared to a similar female who has not tried cannabis. For males, the corresponding school leaving rate is 23 percent. Surprisingly, they find a negative correlation between unobserved characteristics determining the hazard rate for starting cannabis use and the hazard rate for leaving formal education. Contrary to common belief, their results suggest that the estimated effect of cannabis use on school leaving is underestimated when unobserved heterogeneity is ignored.

McCaffrey et al. (2010) aim to establish the causal relationship between marijuana use and education using propensity score methods to reduce the impact of selection bias. They use data that were administered to 61 middle schools in South Dakota in 1997 as part of a large-scale experiment. The baseline survey was administered to 5,857 students and it collected detailed information on individual and family backgrounds. Following the baseline interview of grade 7 students, students were re-interviewed each year until grade 11. Their data show a strong, positive relationship between persistent marijuana use in grades 9 and 10 and the probability of dropping out of high school. However, when they control for background information, academic performance, frequency of cigarette use and selection bias, the relationship is much smaller (an odds ratio of 1.2) and not statistically significant.

Verweij et al. (2013) use a sample of adult twins from the Australian Twin Registry, interviewed between 2006 and 2009, to analyze the effect of early cannabis use (before age 18) on early school leaving (acquire 11 years or less). The use of twin data enables them to examine the potential sources for the association between early cannabis use and educational attainment. Their results suggest that the relationship is not causal but due to overlapping shared environmental influences. They argue that genetic factors may also play a role. A concern with the data used in this study is the reliance on retrospective questions about early marijuana use to respondents who were between 27 and 40 years old at the interview.

Another study designed to assess the possibility of a causal effect of early cannabis initiation and educational attainment is Melchior et al. (2017). Their French data on a sample of respondents who were 22-35 years of age combined with a separate parent study. Early use (before age 17) is contrasted with late use (after age 16) and non-use and the data source provides a rich set of information on both juveniles and parents. Their results indicate that initiating cannabis at an early age causally reduces the probability of graduating from high school and this effect is stronger for female students.

In conclusion, the studies discussed above all show a correlation between cannabis use and low educational achievement using different data sources. Many use logistic regressions, controlling for rich sets of observable characteristics. The results based on linear IV regressions arguably suffer from weak and invalid instruments. The Van Ours and Williams (2009) paper estimates a more appropriate model, more in line with how the data was generated, and they find evidence of significant effects. Unfortunately, their main outcome variable, age when leaving school, is not available in their data and they need to rely on an approximation. The McCaffrey et al. (2010) is a careful analysis study but is limited to a specific state that differs in some important aspects from the national averages. The nature of the data also prevents them from considering the impact of using marijuana at earlier ages, before grade 9.

Our paper addresses many of the shortcomings in the previous literature. Like Van Ours and Williams (2009), we develop and estimate a bi-variate duration model allowing for correlated unobserved heterogeneity. But unlike them, we have detailed, longitudinal information on grade transitions in school and are able to control for a rich set of observed individual and family characteristics. And unlike McCaffrey et al. (2010), we use data from a nationally representative longitudinal study that allows us to estimate the effects of early marijuana use on educational attainment from age 16 and onwards. Unlike any of the papers described above, we estimate the impacts at different grade levels and find that they vary significantly. Moreover, while previous studies have limited attention to an average (across individuals) effect of marijuana use, we demonstrate the existence of substantial heterogeneity in effects across different groups of individuals and the importance of accounting for unobserved heterogeneity.

2.3 Data

In this paper, we utilize data from the 1997 cohort of the National Longitudinal Survey of Youth (NLSY97), which is a nationally representative sample of five cohorts of males and females who were born between 1980 and 1984. The initial interview took place in 1997 and follow-up interviews were conducted annually until 2011 after which it became a biannual survey. NLSY97 gathers information in an event history format, in which dates are collected for the beginning and end of significant life events. In addition, there is detailed information on family background and income as well as on individual scholastic ability.

In our analysis, we remove individuals who were not part of the representative cross-sectional sample in 1997 (this removes oversamples of Blacks and Hispanics). We also excluded individuals who did not provide valid information on the following: family income (at any point between 1997 and 2001), mother's age at birth, family situation at the time of the survey (divorced parents or not), area of residence, number of siblings, mother's education and Armed Forces Qualification Test (AFQT) scores.[1] We exclude those with missing information on any of these variables since they are included as

[1] AFQT scores consists of four components of the Armed Forces Vocational Aptitude Battery (ASVAB): Arithmetic Reasoning (AR), Mathematics Knowledge (MK), Word Knowledge (WK), and Paragraph Comprehension (PC). These scores have been used extensively in research on education using NLSY data. In this paper, we follow Belzil and Hansen (2020) and regress the scores on age and education, in order to adjust for age and educational differences at the time of the test, and use the standardized residual from that regression as the measure of cognitive ability.

covariates in all transition probabilities.[2] Finally, we remove respondents who did not provide any answers to questions related to substance use. After these selections, the sample consists of 2,935 individuals.[3]

We use the information on family income for each individual at ages 16 and 17, if available, and construct an average income measure. If income is only available for one of the years, the average income is replaced by that income. If no income information is available for these ages, we consider the income at earlier ages if available in order to minimize the number of individuals dropped because of missing income. We express income in the year 2000 dollars using the CPI for all urban consumers.

For each individual, we measure schooling attainment using the highest grade completed at a given age and do so from age 16 until age 25. To derive measures of marijuana use, we compile information from questions like 1) have you ever used marijuana?; 2) when did you start using marijuana? and 3) did you use marijuana during the year before the interview? From the responses to these questions, we create individual spells of episodes with marijuana use (and non-use). In the most general model specifications, we also include the use of alcohol and cigarettes and information on these substances was obtained in the same way as information on marijuana use. We code substance use from age 11 onwards.

In Table 2.1 we present the proportions of the sample that has ever used marijuana, cigarettes and alcohol, by age starting at age 11. Close to 14 percent of our sample had used marijuana by age 14 and over 30 percent had used cigarettes and alcohol. Two years later, at age 16, the proportion who has ever used marijuana more than doubled to 29 percent. Use of alcohol and cigarettes also increased substantially. The figures presented in Table 1 match proportions from other data sources and show that, in some cases, substance use starts at very young ages.

[2]These variables are commonly included in the empirical analysis of substance use and education. We decided not to include the father's education in the list mainly because of the large number of missing values for this variable and the skewness in responses to questions about this across the sample (there is a higher fraction of missing among non-white respondents).

[3]These sample selections are similar to those used in previous work on education using NLSY data, such as Keane and Wolpin (1997) and Belzil and Hansen (2020).

Table 2.1: Proportion of respondents that have ever used marijuana, cigarettes and alcohol, by age.

Age	Marijuana	Cigarette	Alcohol	Number of individuals
11	0.005	0.044	0.033	2,935
12	0.025	0.131	0.113	2,935
13	0.059	0.221	0.21	2,935
14	0.138	0.328	0.348	2,935
15	0.221	0.407	0.512	2,935
16	0.293	0.473	0.617	2,935
17	0.379	0.536	0.717	2,875
18	0.463	0.603	0.81	2,799
19	0.524	0.65	0.862	2,720
20	0.569	0.678	0.896	2,643
21	0.597	0.697	0.93	2,578
22	0.619	0.715	0.948	2,506
23	0.637	0.73	0.955	2,434
24	0.654	0.741	0.961	2,370
25	0.661	0.751	0.969	1,870

Details on the highest grade completed and grade transitions by age are presented in Table 2.2. In the first column, we list the average highest grade completed by age which increases from 10.5 at age 16 to just below 14 at age 25. The second column shows the proportion of the sample with a grade increment for each age, starting at age 17. As expected, this proportion drops non-linearly with age with larger drops following normal ages of high school and college graduation (age 18 and 22, respectively).

Table 2.2: Highest grade completed and grade transitions

Age	Highest Grade completed	Grade transition	Number of individuals
16	10.5	-	2,935
17	11.3	0.832	2,875
18	11.9	0.642	2,799
19	12.4	0.496	2,720
20	12.9	0.418	2,643
21	13.3	0.394	2,578
22	13.6	0.253	2,506
23	13.7	0.177	2,434
24	13.8	0.138	2,370
25	13.9	0.114	1,870

In Table 2.3, we show average characteristics separately for individuals who never used marijuana and for those who used it at least once over the sample period. There is no significant gender difference in usage while the table entries suggest that Blacks are somewhat over-represented among the never-users. The proportion living with both biological parents at the interview date is higher among the never-users (0.66) than among the users (0.57). For other background variables - family income, mother's education, AFQT scores, mother's age at birth, urban residence and number of siblings - there are no major differences in sample means between the two groups. There is however a large difference in ever-used cigarettes. Substantially fewer never-use (marijuana) persons have ever used cigarettes, 44 percent versus 90 percent for those who have used marijuana at least once. Lastly, 57 percent of our sample have used marijuana at least once. This figure is similar to the 63 percent reported by Williams and Van Ours (2020) in their paper on cannabis use and school-to-work transitions using NLSY97.

Table 2.3: Sample means, by marijuana use

	Never used	Used at least once
Male	0.47	0.5
Black	0.16	0.12
Hispanic	0.11	0.12
Intact family	0.66	0.57
Family income*	$\$64,273$	$\$67,163$
Mother - high school graduate	0.37	0.34
Mother - attend college	0.48	0.53
AFQT	0.07	0.08
Mother's age at birth	25.9	26.1
Urban	0.7	0.76
Number of siblings	2.4	2.2
Ever used cigarettes	0.44	0.9
Ever used alcohol	0.84	0.99
Drop out of high school	0.18	0.2
High school graduate	0.22	0.22
Some college	0.08	0.08
College graduate	0.33	0.31
Number of observations	1,262	1,673

- * Family income is expressed in year 2000 dollars.

Similar to earlier studies on substance use that utilize retrospective information, our measures of marijuana, cigarettes and alcohol are subject to potential measurement error problems, specifically recall errors. However, unlike Van Ours and Williams (2009) whose sample consists of respondents aged 25-50, the respondents were asked about their substance use at a young age (in some cases at age 12). We, therefore, believe the issue of recall errors is less serious in our paper than in many of the previous studies on this topic.

2.4 Estimation

In this paper, we explore the possibilities of adverse consequences of using marijuana on educational attainment. Exploiting the longitudinal nature of the NLSY97 data, we analyze the timing of marijuana use and school transitions within a duration model framework. We allow for potentially correlated unobservable components of these transitions. As discussed in Van Ours and Williams (2009), a substantial advantage of using a duration approach is that identification of the effect of marijuana use does not rely on a conditional independence assumption like in more commonly used linear, cross-sectional models. Instead, as shown by Abbring and Van Den Berg (2003), treatment effects can be identified from spell data without the need to rely on exclusion restrictions. This result has been used extensively in the previous literature on unemployment duration (see for example Van Den Berg et al. (2004) and Abbring et al. (2005)) and in cannabis use (see Van Ours and Williams (2009)).

Similar to Van Ours and Williams (2009), we assume that the rate at which individuals start using cannabis depends on age, their observed family and individual characteristics as well as on their unobserved characteristics. Individuals are assumed to be at risk of starting to use marijuana as of age 10. Below we detail the specifications of the transition rates and contributions to the likelihood functions for different assumptions on the distribution of unobserved heterogeneity.

The transition rate into cannabis use at age t conditional on observed characteristics X and unobserved characteristics θ is specified as:

$$\lambda_i^m \left(t|\theta_i^m\right) = \frac{exp\left(y_i^m\left(t\right)\right)}{1 + exp\left(y_i^m\left(t\right)\right)}, \tag{2.1}$$

where

$$y_i^m\left(t\right) = X_i\beta^m + \alpha_1^m \ln\left(t\right) + \alpha_2^m \ln\left(t\right)^2 + \alpha_3^m \ln\left(t\right)^3 + \theta_i^m. \tag{2.2}$$

The vector X includes time-invariant personal characteristics. Most are predetermined at the initial age (such as gender, race, mother's education, number of siblings and mother's age at birth) while others are measured at the initial survey in 1997 (such as family stability, parental income, cognitive test score and urban residency). To capture the effects of duration dependence, we follow Ham and Lalonde (1996) and use polynomials of log duration. Unobserved heterogeneity reflects differences in the susceptibility to the uptake of cannabis and is represented by θ_i^m. We provide details on the specifications of unobserved heterogeneity below.

We model educational grade progression as a discrete-time process following Belzil and Hansen (2020). Specifically, we assume that the decision process for education starts at age 16. The choice variable is denoted d_e^t, where $d_e^t = 1$ when an individual invests in an additional grade attainment in period t, $d_e^t = 0$ otherwise. The probability of advancing a grade level in period t is defined as:

$$Pr\left(d_e^t = 1\right) = \frac{exp\left(y_i^e\left(t\right)\right)}{1 + exp\left(y_i^e\left(t\right)\right)}, \tag{2.3}$$

where

$$\begin{aligned} y_i^e\left(t\right) =& X_i\beta^e + \gamma_0 G_i^{15} + \gamma_1 G_{i,t}^{hs} + \gamma_2 G_{i,t}^{ac} + \gamma_3 G_{i,t}^{cg} + \\ & \delta_1^m I_{i,t}^m + \delta_2^m I_{i,t}^m \times G_{i,t}^{hs} + \delta_3^m I_{i,t}^m \times G_{i,t}^{ac} + \delta_4^m I_{i,t}^m \times G_{i,t}^{cg} + \theta_i^e. \end{aligned} \tag{2.4}$$

The observable characteristics, X, are the same as those used for the transition into marijuana. The initial condition, grade level completed at age 15, is represented by G_i^{15}. We allow the grade progression probabilities to depend on completed grade levels. Specifically, $G_{i,t}^{hs}$ equals one if the person has completed high school in period t, and it equals zero otherwise. Similar variables are created for attending college, $G_{i,t}^{ac}$, and graduated from a four-year college program, $G_{i,t}^{cg}$. The δ parameters capture the effects of marijuana use on grade progression since $I_{i,t}^m$ is an indicator variable that equals one if the person started to use marijuana prior to the current period and equals zero otherwise. We further allow the impact of marijuana use to vary across completed grade levels. Unobserved heterogeneity is represented by θ_i^e.

2.4.1 Unobserved Heterogeneity

As mentioned above, and as has been documented in previous studies, controlling for unobserved heterogeneity is critical when analyzing labor market transitions. Ignoring potentially correlated unobserved heterogeneity terms (and self-selection) may result in biased and inconsistent estimates of the transition probabilities. In this paper, we consider alternative specifications for unobserved heterogeneity (θ). In the simplest and most naive specification, each of the two θ terms is represented by a scalar parameter and there is no dependence across marijuana use and education. That is:

$$\theta_k^j = \mu^j, \qquad j = m, e;\ k = 1, 2, \tag{2.5}$$

where μ^j is a fixed intercept. In an alternative and more flexible specification, we assume that each of the two θ terms is discretely distributed with two points of support. For the cannabis spell, the two values of θ_k denote high versus low propensity to initiate consumption of the drug. For grade transitions, they represent high versus low probabilities of accumulating education.

We allow for dependence across spells and this yields four combinations

$$\{(\theta_1^m, \theta_1^e), (\theta_1^m, \theta_2^e), (\theta_2^m, \theta_1^e), (\theta_2^m, \theta_2^e)\}. \tag{2.6}$$

For each combination r, there is an associated probability, P_r:

$$\begin{aligned} P_1 &= Pr(\theta^m = \theta_1^m, \theta^e = \theta_1^e) \\ P_2 &= Pr(\theta^m = \theta_1^m, \theta^e = \theta_2^e) \\ P_3 &= Pr(\theta^m = \theta_2^m, \theta^e = \theta_1^e) \\ P_4 &= Pr(\theta^m = \theta_2^m, \theta^e = \theta_2^e) \end{aligned} \tag{2.7}$$

The type probabilities are estimated assuming a multinomial logit specification:

$$P_r = \frac{exp\left(\eta_r\right)}{\sum_{w=1}^{4} exp\left(\eta_w\right)}, \tag{2.8}$$

where η_4 is normalized to zero.

2.4.2 Likelihood Functions

For the simplest model specification with the scalar representation of unobserved heterogeneity, and for uncensored marijuana non-use spells, the individual contribution to the likelihood function is given by

$$\begin{aligned} L_i = \lambda_i^m\left(Tm|\theta^m\right) \prod_{tm=1}^{Tm-1} \left(1-\lambda_i^m\left(tm|\theta^m\right)\right) \times \\ \prod_{te=16}^{25} Pr\left(d_{e,te}=1|\theta^e\right)^{d_{e,te}} \left(1-Pr\left(d_{e,te}=1|\theta^e\right)\right)^{(1-d_{e,te})}. \end{aligned} \tag{2.9}$$

Not all individuals start using cannabis during the sample period (until age 25 or period 17) and for these individuals, we do not observe the end of the (right censored) spell. In these cases, the contribution to the likelihood is instead given by

$$\prod_{tm=1}^{17} \left(1-\lambda_i^m\left(tm|\theta^m\right)\right). \tag{2.10}$$

For the specification with correlated unobserved heterogeneity, the individual, unconditional contribution to the likelihood function is a weighted average of the likelihood function above using the type probabilities (π_r) as weights.

$$L_i = \sum_{r=1}^{4} \pi_r \times f_m\left(t_m|\theta_r^m\right) \times g\left(d_e|\theta_r^e\right), \tag{2.11}$$

where

$$f_m\left(t_m|\theta_k^m\right) = \lambda_i^m\left(Tm|\theta_k^m\right)\prod_{tm=1}^{Tm-1}\left(1-\lambda_i^m\left(tm|\theta_k^m\right)\right), \qquad k=1,2 \tag{2.12}$$

if uncensored and

$$f_m\left(t_m|\theta_k^m\right) = \prod_{tm=1}^{17}\left(1-\lambda_i^m\left(tm|\theta_k^m\right)\right), \qquad k=1,2 \tag{2.13}$$

if censored. For grade progression,

$$g\left(d_e|\theta_k^e\right) = \prod_{te=16}^{25} Pr\left(d_{e,te}=1|\theta_k^e\right)^{d_{e,te}}\left(1-Pr\left(d_{e,te}=1|\theta_k^e\right)\right)^{\left(1-d_{e,te}\right)}, \qquad k=1,2. \tag{2.14}$$

The likelihood of the sample data is formed by the product of each individual contribution (L_i).

2.5 Empirical Results

In this section, the focus of our discussion is on the marginal effects of selected variables. A full set of parameter estimates and their asymptotic standard errors for each model specification are available in Appendix C. We use a parametric bootstrap to estimate the standard errors of the marginal effects. Specifically, we draw 1,000 vectors of parameter values for the model from the estimated variance-covariance matrix. For each vector, we calculate marginal effects. The reported marginal effects below are the average effects across the 1,000 draws and the standard errors of the effects are estimated using the standard deviation of the simulated effects.

2.5.1 Education

Marginal effects of observable characteristics on grade transitions are presented in Table 2.4. The entries in column one refer to a model specification where marijuana use is assumed to be exogenous. In column two, we model both marijuana and grade transitions with correlated unobserved heterogeneity. Finally, in column three we add information on cigarette and alcohol use and interact these with grade completed at age a. The effects were calculated for each individual and time period and then averaged over time and individuals. A star after the entry indicates the statistical significance of the effect at the common five percent level.

Across all model specifications, most marginal effects are statistically significant and have the expected signs. The only variable for which the marginal effect is not significant for any model specification is urban residency. However, this variable has a significant impact on marijuana use as described below. Overall, females are more likely to advance in school than males and the same is true for Blacks. The intact family has a large positive effect, as does the mother's education. Students from higher-income households progress further in school although the magnitude of the effect is small.[4] As expected, and as documented in previous research on educational attainment, the effect of AFQT is positive and large.

[4]This finding is consistent with Belzil and Hansen (2020) who reports a smaller effect of family income on educational attainment for respondents from the 1997 cohort of NLSY than respondents from the 1979 cohort.

Table 2.4: Marginal effects on grade transitions

	Model 1	Model 2	Model 3
Male	-0.032^*	-0.028^*	-0.022^*
Black	0.037^*	0.048^*	0.032^*
Hispanic	0.008	0.022^*	0.018^*
Intact family	0.063^*	0.055^*	0.044^*
Family income	0.004^*	0.004^*	0.004^*
Mother - high school	0.049^*	0.067^*	0.079^*
Mother - college	0.111^*	0.104^*	0.116^*
AFQT	0.097^*	0.096^*	0.092^*
Mother's age at birth	0.004^*	0.005^*	0.005^*
Urban	-0.005	-0.002	-0.01
Number of siblings	-0.007^*	-0.012^*	-0.008^*
High school graduate	-0.331^*	-0.318^*	-0.386^*
Attend college	-0.242^*	-0.343^*	-0.383^*
College graduate	-0.512^*	-0.671^*	-0.750^*
Initial grade level	0.060^*	0.080^*	0.065^*
Marijuana use, by grade level			
less than high school	-0.237^*	-0.156^*	-0.096^*
	[0.006]	[0.009]	[0.009]
high school graduate	-0.123^*	-0.080^*	-0.041^*
	[0.009]	[0.011]	[0.011]
some college	-0.081^*	-0.055^*	-0.023^*
	[0.007]	[0.011]	[0.010]
college graduate	-0.038^*	-0.025^*	-0.015
	[0.014]	[0.011]	[0.011]
Correlation $[\theta_i^m, \theta_i^e]$	–	-0.135	-0.102

- Standard errors are in brackets.
- (*) signifies statistical significance at the 5 percent level.
- Model 1 refers to a model specification where marijuana use is assumed to be exogenous.
- In Model 2, both marijuana and grade transitions are modeled with correlated unobserved heterogeneity.
- In Model 3, we add information on cigarette and alcohol use and interact these with grade completed at age a.
- A parametric bootstrap was used to estimate the standard errors of the marginal effects.
- The full set of parameter estimate, standard errors and marginal effects are provided in Appendix C.

The bottom panel of Table 2.4 documents the marginal effects of initiating marijuana use at different grade levels. The first row displays the marginal effects of starting to use marijuana before completing grade 12. The effect in column one, -0.237, indicates that the probability of advancing one more grade while in high school is reduced by 23.7 percentage points. This is a very large impact which would offset a two-standard deviation increase in AFQT scores. The corresponding marginal effect for students who have completed high school and enters college is significantly lower, -0.123. For grade increments while in college, the impact is further reduced to -0.081. However, this is still a very large effect compared to the estimated effect of the other included background variables. These effects are estimated after controlling for a relatively rich set of observable characteristics but they do not control for possible self-selection into marijuana use that is linked to unobserved characteristics. It is therefore possible that these effects are incorrectly estimated and exaggerate the negative effects of marijuana use. The more general specifications (Model 2 and Model 3) address this.

Model 2 refers to a specification where we jointly model grade transitions and initiating marijuana use as described in Section 4. The marginal effects of the background variables are generally similar to those reported in column one but this is not the case for the effects of using marijuana. For all grade levels, the estimated effects are reduced. For example, the impact of starting to use marijuana in high school drops from -0.237 to -0.156. However, the negative effects at all grade levels remain statistically significant at the five percent level.

We estimate two support or mass points for each source of unobserved heterogeneity and the proportion with $\theta^m = \theta_1^m$, the type with a lower incidence of initiating marijuana, is 0.441 and the proportion with $\theta^m = \theta_2^m$ is 0.559. For education, 52 percent have $\theta^e = \theta_1^e$ (a lower probability of a grade transition) and 48 percent have $\theta^e = \theta_2^e$. The most common combination of θ^m and θ^e (32.6 percent) is $\theta^m = \theta_2^m$ and $\theta^e = \theta_1^e$, signifying a higher probability of marijuana initiation combined with a lower probability of a grade transition. The least common combination (at 19.7 percent) is $\theta^m = \theta_1^m$ and $\theta^e = \theta_1^e$, that is, a lower probability of marijuana initiation and a lower probability of a grade transition. The correlation between θ^m and θ^e is -0.135. This is an expected sign and the magnitude is non-negligible. It further indicates the existence of self-selection into marijuana use and that models that ignore this will overestimate the effect of marijuana on educational attainment. This explains why a reduction in the effects of marijuana is observed when comparing effects from Model 2 with those from Model 1.[5]

The entries in the third column, Model 3, are derived from a specification similar to Model 2 but with the addition of controls for cigarette and alcohol consumption. Specifically, these two variables are derived in the same way as marijuana use and, like marijuana, are interacted with completed grade levels to allow for differential impacts across completed schooling levels. McCaffrey et al. (2010) show that adding controls for cigarette consumption significantly reduce the effect of marijuana use in their model of high school dropouts. Similar to their results, we find that the marginal effect of marijuana use is further reduced when we add cigarette and alcohol consumption. The effect of using marijuana before grade 12 equals -0.096, down from -0.156 obtained in an identical model apart from the cigarette and alcohol covariates. The effects for other grade levels are similarly reduced but remain statistically significant (except for the effect on grade transitions as a college graduate). As discussed by McCaffrey et al. (2010), the use of cigarettes appears to be significantly correlated with some unobserved characteristics that contribute to the observed association between marijuana use and educational attainment.[6] Lastly, the correlation between θ^m and θ^e is -0.102, slightly lower than the correlation obtained for Model 2.[7]

2.5.2 Transitions into Marijuana Use

Marginal effects of observable characteristics on transitions into marijuana use are presented in Table 2.5. The entries suggest that there is no significant gender difference in these transitions. The effect of Black is large, negative and significant indicating that Blacks are significantly less likely to start using marijuana than other racial groups

[5] This contrasts the result in Van Ours and Williams (2009) who surprisingly finds evidence for the opposite relationship.

[6] McCaffrey et al. (2010) examine a number of potential variables that may explain the impact of adding cigarette consumption to their regression equation and find that peer effect is one such variable.

[7] The estimated proportions of θ^m and θ^e for the four combinations in this specification are: $Pr\left(\theta_1^m, \theta_1^e\right) = 0.207$; $Pr\left(\theta_1^m, \theta_2^e\right) = 0.219$; $Pr\left(\theta_2^m, \theta_1^e\right) = 0.338$; and $Pr\left(\theta_2^m, \theta_2^e\right) = 0.236$.

(white and Hispanics) after controlling for a selection of observed characteristics and unobserved heterogeneity. The transition probability is 4.3 percentage points lower for Blacks than whites. A similarly sized and signed marginal effect is found for living in an intact family at the initial interview date in 1997.[8] Family income and mother's education are not significantly related to marijuana use but there is a significant, negative effect of cognitive ability (represented by standardized values of AFQT scores). The estimated effect of -0.011 suggest that a one standard deviation increase of the test score reduces the probability of starting to use marijuana with 0.011. The age of the respondent's mother when he or she was born is negatively related to marijuana use while the opposite is true for those living in urban areas. Finally, the transition probability is significantly lower in larger families.

Table 2.5: Marginal effects on transition into marijuana use

	Model 3
Male	0.003
Black	-0.043^*
Hispanic	-0.005
Intact family	$-0.033*$
Family income	0.001
Mother - high school	0.001
Mother - college	0.009
AFQT	-0.011^*
Mother's age at birth	-0.002^*
Urban	0.022^*
Number of siblings	-0.012^*

- * signifies statistical significance at the 5 percent level.
- The marginal effects are obtained from Model 3 described in Table 2.4.
- A parametric bootstrap was used to estimate the standard errors of the marginal effects.
- The full set of parameter estimates, standard errors and marginal effects are provided in Appendix C.

[8]The variable intact family equals one if the respondent resided with both biological parents in 1997, zero if not.

2.5.3 Age at Marijuana Initiation

In Table 2.6 we present marginal effects obtained from an adjusted version of our preferred model specification where we allow the grade-specific effects of marijuana use to depend on the age when the person first used marijuana. Specifically, we estimate the effects separately for those who started when they were 14 or younger and for those who started when they were 15 or older. There are two main findings from this extended model. First, the negative impact of marijuana use is stronger for those who started at a young age (the effects in high school are -0.112 for early users versus -0.079 for late users). This is true for all grade levels.

Secondly, the negative effect from marijuana use is persistent across grade levels for those who started at a young age. There is only a marginal reduction for high school graduates (-0.106 versus -0.112) and the negative effect is substantial even at college (-0.09). For those who started after age 14, the negative effects drop with completed grade levels and are not significant for grade levels above high school.

Table 2.6: Marginal effects on grade transitions, by age at marijuana initiation

Marijuana use, by grade level	Initiate marijuana at age 14 or younger	Initiate marijuana at age 15 or older
less than high school	-0.112^*	-0.079^*
	[0.011]	[0.007]
high school graduate	-0.106^*	-0.023^*
	[0.012]	[0.01]
some college	-0.090^*	-0.011
	[0.012]	[0.01]
college graduate	-0.044^*	-0.005
	[0.014]	[0.009]

- Standard errors in Brackets.
- * signifies statistical significance at the 5 percent level.
- A parametric bootstrap was used to estimate the standard errors of the marginal effects.
- The full set of parameter estimates and marginal effects are provided in Appendix C.

2.5.4 Heterogeneous Effects

The marginal effects presented so far have all been averaged across individuals in the sample. However, it is reasonable to expect that marijuana use will impact different (groups of) individuals differently. For example, Van Ours and Williams (2009) show significantly larger impacts of marijuana use on school leaving rates for females than for males. In this section, we will explore differences in marginal effects across gender, race, family income, intact family status and mother's age when the respondent was born. We limit the presentation and discussion to the effects of marijuana use on grade transitions. Full set of estimated parameters and corresponding marginal effects are provided in Appendix C.

In Table 2.7, the effects for males appear in column one and those for females are shown in column two. The effects were obtained after estimating Model 3 separately for males and females. There is a significant negative effect of marijuana use on grade transitions in high school and the effect drops as we consider grade transitions after high school. For none of the groups are the effects statistically significant for transitions in college. Unlike Van Ours and Williams (2009), our results show a larger impact of marijuana use on school leaving for males than females. Specifically, the effect of marijuana use on grade transitions in high school is -0.116 for males compared to -0.074 for females, almost 40 percent lower.

Table 2.7: Marginal effects on grade transitions, by gender

Marijuana use, by grade level	Males	Females
less than high school	-0.116^{*}	-0.074^{*}
	[0.013]	[0.015]
high school graduate	-0.040^{*}	-0.045^{*}
	[0.014]	[0.016]
some college	-0.024	-0.019
	[0.015]	[0.015]
college graduate	-0.011	-0.013
	[0.017]	[0.017]
Correlation (θ_i^m, θ_i^e)	-0.044	-0.154

- Standard errors in Brackets.
- * signifies statistical significance at the 5 percent level.
- A parametric bootstrap was used to estimate the standard errors of the marginal effects.
- The full set of parameter estimates and marginal effects are provided in Appendix C.

In Table 2.8, we explore differences in effects across three racial groups: White, Black and Hispanic.[9] For the former two groups, the estimated effects are similar and decline with grade levels. For Hispanic youth, the estimated effect of starting to use marijuana in high school is -0.073, lower than the -0.1 estimated for white and Black respondents. However, the effect associated with starting to use marijuana as a high school graduate is larger for Hispanics (-0.109) than for the other two groups (-0.025 for whites and -0.048 for Blacks).

Table 2.8: Marginal effects on grade transitions, by race

Marijuana use, by grade level	White	blacks	Hispanics
less than high school	-0.100^{*}	-0.100^{*}	-0.073^{*}
	[0.009]	[0.023]	[0.016]
high school graduate	-0.025^{*}	-0.048^{*}	-0.109^{*}
	[0.011]	[0.021]	[0.02]
some college	-0.015	-0.018	-0.016
	[0.009]	[0.02]	[0.02]

- Standard errors in Brackets.
- * signifies statistical significance at the 5 percent level.
- A parametric bootstrap was used to estimate the standard errors of the marginal effects.
- The full set of parameter estimates and marginal effects are provided in Appendix C.

[9] The effects were estimated by interacting the indicator for marijuana use in period t with indicators for Black and Hispanic.

We analyze additional heterogeneity in effects in Table 2.9.The top panel shows marginal effects when we split the sample into poor and non-poor, based on reported family income. Those whose family income is in the lowest quartile are classified as poor, the rest are considered non-poor. The estimated negative effects are stronger for poor youth at all grade levels and also persistent across grade levels. The effects are -0.108 for those who started using marijuana in high school and -0.097 for those who started using marijuana in college. For the non-poor, there is a relatively large negative effect for high school transitions (-0.082) but not for higher grade levels.

Table 2.9: Marginal effects on grade transitions, by group

Marijuana use, by grade level	Poor	Not poor
less than high school	-0.108^{*}	-0.082^{*}
	[0.010]	[0.008]
high school graduate	-0.115^{*}	-0.030^{*}
	[0.013]	[0.008]
some college	-0.097^{*}	-0.02
	[0.015]	[0.011]
	Single parent	Intact family
less than high school	-0.093^{*}	-0.076^{*}
	[0.010]	[0.010]
high school graduate	-0.068^{*}	-0.028^{*}
	[0.014]	[0.012]
some college	-0.040^{*}	-0.016
	[0.015]	[0.009]
	Teen mother	Non-teen mother
less than high school	-0.119^{*}	-0.083^{*}
	[0.010]	[0.008]
high school graduate	-0.068^{*}	-0.030^{*}
	[0.014]	[0.007]
some college	-0.040^{*}	-0.022^{*}
	[0.015]	[0.009]

- Standard errors in Brackets.
- * signifies statistical significance at the 5 percent level.
- A parametric bootstrap was used to estimate the standard errors of the marginal effects.
- Those whose family income is in the lowest quartile are classified as poor.
- Single versus nuclear refers to family status at the base interview in 1997.
- Teen equals one if the respondent's mother was a teenager when they were born, zero otherwise.
- The full set of parameter estimates and marginal effects are provided in Appendix C.

In the second panel of Table 2.9, we show effects when we instead split the sample based on intact family status in 1997. As shown in Table 2.3, 66 percent of the 'never used' lived with both biological parents in 1997 (i.e. intact family), while for 'ever used' the proportion is lower, 57 percent. The effects in panel two of Table 2.9 show that marijuana use has a larger negative impact on those who lived with only one of their biological parents. Again, the effect declines with grade level but for all grade levels, the negative effects are stronger for this group than for those who lived in a nuclear family.

Finally, in the bottom panel of Table 2.9 we present the marginal effects when we allow the effects to vary depending on the mother's age at birth. Specifically, the entries in column one show the effects for respondents whose mother was a teenager when they were born. The effects in column two are for respondents with non-teen mothers. The effects are more pronounced for youth to teenage mothers and they do not decline as fast as for the other group.

2.5.5 Peer Effects

McCaffrey et al. (2010) showed a significant impact on the effect of marijuana use on high school dropout decisions when they added a control for peer effects. To explore if our results are similarly sensitive towards the inclusion of peer effects, we constructed a measure from nine survey questions on peers and added it as a covariate in our preferred model specification. Four of the nine questions ask respondents about the percentage of their peers that engage in activities that are likely to be positively related to education: (i) attend church regularly; (ii) participate in sports, clubs or school activities; (iii) plan to go to college; and (iv) do volunteer work. The remaining five questions are related to activities that are likely to be negatively related to education: (i) smoke cigarettes; (ii) get drunk at least once per month; (iii) use illegal drugs; (iv) skip classes; and (v) belong to a gang. For each question, the response options are: less than 10 percent; about 25 percent; about 50 percent; about 75 percent; and more than 90 percent.

We use responses to each of the nine questions and combine them into a single measure of peers by taking the average for each individual, inverting the answers to the five questions that seek to identify the presence of poor peer influence. Thus, for our measure, higher values of the peer average are associated with better peers. Finally, we standardize the measure to have mean zero and variance one.

Table 2.10: Marginal effects on grade transitions, with and without peer effects

Marijuana use, by grade level	Excluding peer effects	Including peer effects
less than high school	-0.096^{*}	-0.094^{*}
	[0.009]	[0.010]
high school graduate	-0.041^{*}	-0.041^{*}
	[0.011]	[0.011]
some college	-0.023^{*}	-0.017
	[0.010]	[0.012]
college graduate	-0.015	-0.012
	[0.011]	[0.012]
Correlation (θ_i^m, θ_i^e)	-0.102	-0.083

- Standard errors in Brackets.
- * signifies statistical significance at the 5 percent level.
- A parametric bootstrap was used to estimate the standard errors of the marginal effects.
- The full set of parameter estimates and marginal effects are provided in Appendix C.

The peer variable is added to both transitions (into marijuana and grade increment) and has the expected negative effect on marijuana initiation and positive effect on grade transitions. In the former case, the mean effect is -0.029, similar to the effect for residing in an intact family in 1997. The effect on grade transitions is 0.027 which is relatively large. However, while peer effects are significant for the transitions, the inclusion of this variable does not change the marginal effects of marijuana use on education. The entries in Table 2.10 are derived from a model specification that is identical to the one that generated the results in column 3 of Table 2.4. That is, we allow for correlated unobserved heterogeneity and include controls for cigarette and alcohol consumption. We repeat the effects from Table 4 in the first column in Table 2.10 (model without peer effects) and then show the corresponding marginal effects when we add peer effects. The effects of marijuana use in high school and as a high school graduate are very similar across the two model specifications. For impacts on transitions beyond high school, they are not statistically significant in the model with peer effects.

2.5.6 Role of Unobserved Heterogeneity

The results presented in Table 2.4 showed the importance of controlling for unobserved heterogeneity in estimating the effects of marijuana use on educational attainment. In this section, we illustrate to what extent the transition probabilities depend on this heterogeneity by calculating predicted transition rates for both marijuana use and grade increments for different types of individuals. To contrast the impact of unobserved heterogeneity, we calculate the transition rates for two groups based on their observed characteristics. We label the two groups as 'at risk' and 'not at risk'. The former group consists of hypothetical individuals: 1) who did not live in an intact family; 2) who had a family income of \$20,000; and 3) whose mother's were 17 when they were born. The entries in Table 2.9 showed that youth with these characteristics are more adversely affected by marijuana use than other groups.

The predicted transition rates are presented in Table 2.11. The first two rows show transition rates into (first) marijuana use at age 15, first for individuals identified by us as being at risk and then for the 'not at risk' group. The top row entries are evaluated when $\theta^m = \theta^m_1$, a lower probability of starting to use marijuana while the entries in row two are evaluated when $\theta^m = \theta^m_2$, a higher probability of starting to use marijuana. For the first group of individuals, the probability of starting to use marijuana at age 15 is low for both at risk and not at risk individuals, 0.024 and 0.011, respectively. For the second group, the probabilities are significantly larger, 0.352 and 0.197, respectively. The difference between the two groups, given risk status, is substantially larger that the difference between at risk and not at risk, with groups. This illustrates the importance of incorporating unobserved heterogeneity in the model.

Table 2.11: Predicted transition rates, by risk group and unobserved type

	Proportion of sample	At risk	Not at risk
Unobserved type	Transitions into marijuana use at age 15:		
$\theta^m = \theta^m_1$	0.426	0.024*	0.011*
		[0.006]	[0.003]
$\theta^m = \theta^m_2$	0.574	0.352*	0.197*
		[0.027]	[0.012]
Unobserved type	Grade transitions:		
$\theta^e = \theta^e_1$	0.544	0.316*	0.486*
		[0.012]	[0.008]
$\theta^e = \theta^e_2$	0.456	0.754*	0.834*
		[0.008]	[0.004]

- Standard errors in Brackets.
- * signifies statistical significance at the 5 percent level.
- A parametric bootstrap was used to estimate the standard errors of the predicted probabilities.
- At risk equals one for hypothetical individuals who:
1) did not live in a nuclear family;
2) had a family income of $20,000;
3) whose mothers' were 17 when they were born.
- Predictions were obtained using estimates from Model 3 in Table 2.4.

The bottom part of Table 2.11 shows grade transition rates at age 17. For both types, ($\theta^e = \theta^e_1$) and ($\theta^e = \theta^e_2$), the transition rates are lower for the 'at risk' group (0.316 for the former group and 0.754 for the latter) compared to the 'not at risk' group (0.486 and 0.834, respectively). Similar to the transitions into marijuana use, the differences across types are larger than the differences across risk groups, conditional on the type.

Finally, in order to further illustrate the importance of unobserved heterogeneity, we present estimated impacts of marijuana use on grade transitions, by grade level, for each of the four possible types (see Equation 2.1). To predict type membership, we use Bayes' theorem and derive the probability that individual i belongs to type k, conditional on observed variables $\mathbf{x}$ as follows:

$$\Pr\left(k|\mathbf{x},\Theta,\pi\right) = \frac{\pi_k f_m\left(t_m|\theta^m_k\right) \times g\left(d_e|\theta^e_k\right)}{\sum_{k=1}^{K} \pi_k f_m\left(t_m|\theta^m_k\right) \times g\left(d_e|\theta^e_k\right)}. \tag{2.15}$$

We then assign individual i to type k if

$$\Pr(k|\mathbf{x}, \Theta, \pi) = max\{\Pr(j|\mathbf{x}, \Theta, \pi)\}, \ j = 1, ..., K. \tag{2.16}$$

The results are shown in Table 2.12. There is evidence of significant heterogeneity in the negative impact of marijuana use on grade progression. The effect for those who started using marijuana in high school is -0.13 for individuals assigned to type one (who have lower probabilities of initiating marijuana and grade transitions) and -0.148 for individuals assigned to type three (higher probability of initiating marijuana and lower probability of grade transitions). The negative effects are substantially smaller for individuals with higher grade transition probabilities, especially when this is combined with a low probability of initiating marijuana (-0.028). This variation highlights the existence of important group-level differences in the effect of marijuana use on school performance. For many students, the negative effect is small, just over 23 percent of the sample is predicted to belong to type 3 (the type with the smallest negative effect). However, for half the sample, start using marijuana in high school reduces grade progression by 13-15 percentage points.

Table 2.12: Marginal effects of marijuana use on grade transitions, by grade level and unobserved type

	Type			
Transition into marijuana	$\theta^m = \theta^m_1$	$\theta^m = \theta^m_1$	$\theta^m = \theta^m_2$	$\theta^m = \theta^m_2$
Grade transition	$\theta^e = \theta^e_1$	$\theta^e = \theta^e_2$	$\theta^e = \theta^e_1$	$\theta^e = \theta^e_2$
Grade level				
less than high school	-0.130^*	-0.028^*	-0.148^*	-0.056^*
	[0.010]	[0.002]	[0.011]	[0.005]
some college	-0.033^*	-0.030^*	-0.025^*	-0.037^*
	[0.012]	[0.010]	[0.008]	[0.013]
Proportion of sample	20.9%	23.2%	33.1%	22.8%

- Standard errors in Brackets.
- * signifies statistical significance at the 5 percent level.
- A parametric bootstrap was used to estimate the standard errors of the predicted probabilities.
- The marginal effects were obtained using estimates from Model 3 in Table 2.4.

As shown earlier, the negative effect of marijuana use is smaller for grade transitions in college than in high school. Further, there is also less variation in this effect across the four types. The entries in the second row of Table 2.12 show negative effects varying from -0.025 to -0.037. None of these effects are significantly different from each other although they are all significantly different from zero.

2.5.7 Intensity of Consumption

The analysis so far has been based on time until first marijuana use, regardless of how much or often the respondent consumed marijuana. While this is the dominant approach in the literature, it seems reasonable that the effect of marijuana consumption on educational attainment depends on the amount or intensity of consumption. To exploit this further, we use questions in the NLSY97 on how many days in the past month (that is, the month preceding the interview) the respondent consumed marijuana. However, this question was first asked in 1997 and we consequently don't have this information before age 12 for anyone.[10] To reduce the issue with initial conditions, we exclude respondents who were 14 or older in 1997 and assume that nobody consumes marijuana before age 13.[11]

In Table 2.13 below, we report the proportion that has ever used marijuana, by age, based on the reported intensity of consumption in the month preceding the interview. The entries in the first column show the proportion, by age, that used marijuana at least one day during the month before the interview took place while those in the second column show the corresponding proportions for those who used it at least 10 days. The proportions for at least one day per month are similar to those reported in Table 2.1 until age 17. After that, the entries in Table 2.13 are somewhat lower than those in Table 2.1 and by age 25, about 54 percent have ever used marijuana.

When we consider more intense use, the proportions are naturally lower but show

[10] The respondents were aged 12-16 in 1997.

[11] Among the 12-year-olds at the time of the survey, only four reported that they used marijuana that year. We have no reason to assume that this is more common among the 13-year-old in 1997. Hence, we believe this assumption is reasonable.

the same increases by age. At age 25, just over 24 percent have ever used marijuana at least 10 days per month.

Table 2.13: Proportion of respondents that have ever used marijuana by age.

	Marijuana use		
Age	At least 1 day per month	At least 10 days per month	Number of individuals
13	0.055	0.013	1,172
14	0.132	0.035	1,172
15	0.225	0.072	1,172
16	0.303	0.104	1,172
17	0.372	0.154	1,133
18	0.422	0.19	1,091
19	0.449	0.208	1,044
20	0.472	0.226	991
21	0.495	0.236	944
22	0.515	0.247	912
23	0.517	0.252	865
24	0.527	0.255	830
25	0.537	0.244	430

In order to determine the impact of intensity of marijuana consumption on education, we use the reduced sample and estimate versions of the model presented above. In particular, we use the most flexible model specification presented in Table 2.4, adding information on peers. However, we exclude information on cigarette consumption and only include marijuana and alcohol use. The reduction in the marginal effects of marijuana use when we added alcohol and cigarette consumption, reported in Table 2.4, were mainly due to the inclusion of alcohol.

The marginal effects are presented in Table 2.14 below. The effect of marijuana use in high school, when any use is considered (-0.105), is similar to the corresponding effect reported in column three of Table 2.4 (-0.096). The effect for high school graduates is however larger in the younger sample (-0.092 versus -0.041). The impact on transitions in college is also larger in the younger sample (-0.049) than in the original sample used to produce Table 2.4 (-0.023).

The effects from intensive consumption (10 days or more per month) are reported in column two of Table 2.14. The magnitudes are similar to, and not statistically significantly different from, those presented in column one except for the case of transitions in college. The effect while in high school (-0.090) is slightly lower than the corresponding effect in column one (-0.092) while the opposite is true for high school graduates. Thus, the estimated marginal effects of marijuana use on grade transitions in high school and as a high school graduated that we have presented in this paper are robust towards different levels of consumption. However, this is not the case for transitions in college where the negative impact of heavy or intense marijuana consumption is similar to the effects on transitions in high school.

Table 2.14: Marginal effects on grade transitions, by intensity level.

	Marijuana use	
Grade level	At least 1 day per month	At least 10 days per month
less than high school	-0.105^{*}	-0.090^{*}
	(0.011)	(0.012)
high school graduate	-0.092^{*}	-0.098^{*}
	(0.018)	(0.020)
some college	-0.049^{*}	-0.097^{*}
	(0.015)	(0.016)
College graduate	-0.001	0.009
	(0.016)	(0.019)
Correlation (θ_i^m, θ_i^e)	0.020	0.041

- Standard errors in Brackets.
- * signifies statistical significance at the 5 percent level.
- A parametric bootstrap was used to estimate the standard errors of the marginal effects.
- The full set of parameter estimates and marginal effects are provided in Appendix C.

2.6 Conclusions

In this paper, we provide further evidence of the impact of using marijuana on educational outcomes. This topic is important for many reasons, one being the fact that educational outcomes are strongly correlated with future labor market success. To understand this relationship between marijuana use and education is perhaps more important than ever given the recent legalization of marijuana use in many jurisdictions. As has been documented in previous research on this topic, there is a clear and significant negative correlation between marijuana use and educational attainment. What is less clear is to what extent that relationship is causal. There is a common understanding that marijuana use is endogenous to the educational investment process and both outcomes are partly determined by common factors. Unfortunately, many of these factors are not observed is representative survey data. A popular approach to address endogeneity in linear regression models is Instrumental Variables (IV). However, many of the instruments that have been utilized for marijuana use lack power and the resulting IV estimates are inconclusive.

Two contributions to this literature, McCaffrey et al. (2010) and Van Ours and Williams (2009) do not rely on IV methods. Instead, McCaffrey et al. (2010) address the selection (on unobservables) issue by adopting a propensity score estimator while Van Ours and Williams (2009) estimate a bi-variate duration model. The McCaffrey et al. (2010) is a careful analysis but is limited to a specific U.S. state (South Dakota) that differs in some important aspects and outcomes from national averages. The nature of the data also prevents them from considering the impact of using marijuana at earlier ages, before grade 9. Van Ours and Williams (2009) estimate a more appropriate model, in line with how the data was generated, and find evidence of significant effects. Unfortunately, their main outcome variable, age when leaving school, is not available in their data and they need to rely on an approximated age. Further, their sample is drawn from a cross-section of respondents aged 25-50 who were asked retrospective questions regarding the age of marijuana initiation, introducing the possibility of bias due to recall errors.

Our paper addresses some important shortcomings in the previous literature. Like Van Ours and Williams (2009), we develop and estimate a bi-variate duration model allowing for correlated unobserved heterogeneity. But unlike them, we have detailed information on grade transitions in school. Moreover, respondents in our sample were asked about substance use at a much younger age (in some cases at age 12) and we argue that the issue of recall errors is less serious in our paper than in theirs. And unlike McCaffrey et al. (2010), we use data from a nationally representative longitudinal study that allows us to estimate the effects of early marijuana use on educational attainment from age 16 and onward. Unlike any of the papers in the related literature, we estimate the impacts of marijuana use at different grade levels and find that they vary significantly. Another important contribution is the analysis of heterogeneity in effects across different individuals and the importance of unobserved heterogeneity.

Our results suggest that marijuana use has a significant, negative effect on grade transitions, both in high school and in college. The negative effect declines with the highest grade completed and is largest in high school. We show that a model without controls for correlated unobserved heterogeneity severely exaggerates the negative effects of marijuana use at all grade levels. Similar to McCaffrey et al. (2010), we find that omitting the consumption of tobacco and alcohol generates larger negative effects. The average marginal effects from our preferred model specification indicate that starting to use marijuana while in high school reduces next year's grade transition by 9.6 percentage points and by 2.3 percentage points while in college. The corresponding effects in the naive, single spell model are 23.7 and 8.1 percentage points, respectively. We estimate larger negative effects for males (11.6 percentage points) than for females (7.4 percentage points) in high school and larger negative effects for Hispanic students than other students. We show that the negative effect of marijuana use is stronger for youth from weaker family backgrounds (low-income, single mothers and teenage mothers) and that unobserved heterogeneity explains a substantial proportion of the variation in marijuana initiation and grade transitions, even after controlling for a rich set of observed characteristics. Finally, we show that the results are robust to different consumption levels. The marginal effects of marijuana use on grade transitions in high school when we define consumption based on any use are similar to those obtained when we record consumption only if the person used it 10 times or more per month.

Similar to previous studies on this topic, a couple of potential limitations need to be considered. The first is the possibility of reversed causality where poor schooling outcomes cause students to start using marijuana. The second is measurement error, especially regarding the reporting of marijuana use. We believe it is difficult to address the first issue with our data. Many respondents in NLSY report starting marijuana use at a young age (over 22 percent started before age 16) and we lack good measures of educational outcomes for many respondents at these young ages. Regarding measurement error in marijuana use, we believe the fact that respondents were asked about their substance use at a young age (in some cases at age 12) reduces this problem.

Chapter 3

Persistent Marijuana Use: Evidence from the NLSY

Abstract

We analyze persistence in marijuana consumption utilizing data from the 1997 cohort of the National Longitudinal Survey of Youth (NLSY97). We allow for three sources of persistence: pure state dependence, time-invariant unobserved heterogeneity and persistence in idiosyncratic, time-varying shocks. We also consider intensity of consumption based on days of use per month and estimate a dynamic ordered Probit model using simulated Maximum Likelihood. We consider a Polya model that generalizes the more commonly used Markov models. The results show that there is a causal effect of previous use. However, ignoring unobserved heterogeneity and serially correlated shocks significantly exaggerates the state dependence.

JEL Code: I12,I21

Keywords: marijuana; persistence; state dependence; unobserved heterogeneity; dynamic ordered probit; simulation; NLSY

3.1 Introduction

The legal status of recreational marijuana in the US has changed significantly since 2012 when Colorado and Washington became the first states to legalize cannabis for adult use. Currently, recreational use is legal in as many as 18 states plus the District of Columbia. These changes have occurred despite evidence pointing to negative impacts from marijuana use (especially at young ages) on different outcomes, such as educational attainment (Hansen and Davalloo, 2022), school to work transitions (Williams and Van Ours, 2020), financial and relational difficulties in adulthood (Chan et al., 2021; Cerdá et al., 2016), health (Hall and Degenhardt, 2009; Lev-Ran et al., 2014), and welfare use and unemployment (Fergusson and Boden, 2008; Schmidt et al., 1998). Marijuana consumption has also been shown to increase the risk of consuming hard drugs (see Deza, 2015).

However, it is possible that the nature of marijuana consumption and its associated risks are heterogeneous in the population. For many, consumption is modest, occasional and highly transitory, while others use marijuana on a regular and persistent basis, and the existence and magnitude of any negative impacts of marijuana use are likely to vary with consumption patterns. However, if there is a causal, addictive effect of marijuana use over time, any initiation is associated with a risk of continued, persistent use. In this case, policies that make marijuana consumption more accessible and socially acceptable may therefore increase the risk of marijuana dependence. On the other hand, if there is no causal effect of past marijuana use on current consumption, this risk is eliminated. Therefore, it is important to understand the dynamics of marijuana consumption and how it varies, at an individual level, over time.

In this paper, we analyze transitions into and out of marijuana consumption. Data from the 1997 cohort of the National Longitudinal Survey of Youth (NLSY97) show that the probability of using marijuana in a given year is almost two times higher for those who used it the year before compared to those who did not use it. However, this data pattern is uninformative about the nature of marijuana persistence. Does past consumption cause current use (perhaps by changing preferences for the drug)? Or is the data simply reflecting different innate propensities to use marijuana over time where some youth receive substantial utility from marijuana consumption and therefore continuously use it while others receive a negative utility and never use it. A third possibility for the observed time dependence is persistence in random shocks to the utility of consumption. For example, an event in school or within the family may alter the perceived utility and induce consumption in a given year. This effect may then persist over time. This paper aims to estimate the sources of persistence in marijuana consumption and evaluate their relative importance for overall persistence.

Our empirical framework builds on the influential work by Heckman (1981) and others who have developed models designed to separate true state dependence from spurious dependence (due to persistent unobserved heterogeneity). These models have been estimated for a number of different outcomes, such as welfare (David Card and Hyslop, 2005; Hansen and Lofstrom, 2009), labor supply (Hyslop, 1999), unemployment (Hansen and Lofstrom, 2009) and health (Carro and Traferri, 2014). A particularly relevant study for this paper is Deza (2015), who uses a dynamic discrete choice model to analyze persistence in illicit drug use. Using data from the 1997 cohort of the NLSY, she estimates a general model of alcohol, marijuana and hard drug use and separates the contributions from state dependence and unobserved heterogeneity, both within drugs and also between drugs. Her results show the existence of significant "stepping-stone" effects into hard drugs, where current alcohol and marijuana use significantly increase the probability of hard drug use in the future.

Our paper addresses some important shortcomings in the previous literature. We first analyze the probability of marijuana use among American youth from ages 13 to 26, paying particular attention to its persistence. Apart from Deza (2015), there are few studies that have analyzed time dependence or persistence in marijuana consumption. While Deza (2015) estimates a general, dynamic model of consumption of alcohol and hard drugs, in addition to marijuana, the focus is on structural state dependence and transitions from alcohol and marijuana into hard drugs (that is, if softer drugs serve as "stepping-stones" into hard drugs). Our model specification, while limited to marijuana consumption only, allows for more general forms of dynamics as well as serially correlated utility shocks. We also estimate different persistence probabilities conditional on the amount consumed, allowing for the separation of occasional or experimental use from continuous, intensive use. We show that these additional dimensions are important and that moderate consumption of marijuana may serve as a "stepping-stone" into heavy use.[1]

The results indicate that serial correlation in the time-varying utility shocks contributes substantially to overall observed persistence. If ignored, the estimate for structural state dependence and the estimated variance of persistent unobserved heterogeneity are exaggerated, leading to incorrect inferences about sources of persistence. Further, separating moderate use from intense use is important.

Focusing first on the estimated average partial effects, which are designed to show the causal effect of past consumption on current consumption, our results for the most general specification of the binary case suggest that consumption of marijuana in the previous period increases the probability of current consumption by 0.129.[2] Given an unconditional consumption rate of 15-20 percent (depending on age), this effect is very large. However, it is still significantly smaller than the corresponding effect obtained from a one-period lagged Markov model (where the effect is 0.192).

We estimate two average partial effects for the ordered model for each intensity level. For moderate consumption levels, the first effect is the difference in conditional probabilities of moderate consumption when we condition on moderate versus no consumption in the previous time period, while the second effect conditions on moderate and heavy use instead. The former effect (moderate versus no consumption) is 0.046, while the second effect is -0.051. That is the probability of consuming moderate levels

[1]We define moderate use as consumption less than nine times per month and heavy use as ten days or more of consumption. The data show that persistence is concentrated among heavy users while moderate use is more transitory. Specifically, the average probability of heavy marijuana use is 0.164, conditional on moderate consumption in the previous time period. This should be compared to a probability of 0.021 among those who did not use marijuana in the previous period.

[2]The average partial effect is estimated as $\hat{Pr}\left(y_{i,t}=1|y_{i,t-1}=1\right)-\hat{Pr}\left(y_{i,t}=1|y_{i,t-1}=0\right)$, which is averaged across individuals and time periods.

of marijuana in year t is 4.6 percentage points higher if the person consumed the same level of marijuana in year $t-1$, relative to not using any marijuana in year $t-1$. While the magnitude of this effect is smaller than the one obtained in the binary case, it constitutes a relative effect close to 50 percent, given the proportions of moderate consumption observed in the data. The negative effect for moderate versus heavy usage suggests a higher probability of moderate use in year t for those with heavy consumption in the previous year compared to those with moderate consumption.

For heavy consumption levels, the first effect is the difference in conditional probabilities of heavy consumption when we condition on heavy versus no consumption in the previous time period, while the second effect conditions on heavy and moderate use instead. The former effect equals 0.043, similar to the one estimated for moderate use. The second effect is smaller, 0.027. That is, the probability of consuming heavy levels of marijuana in year t is 4.3 percentage points higher if the person consumed the same level of marijuana in year $t-1$, relative to not using any marijuana in year $t-1$. Again, while the magnitude of this effect is smaller than the one obtained in the binary case, it constitutes a relative effect close to 50 percent, given the moderate consumption rates observed in the data.

Finally, our analysis of the sources for persistence in marijuana consumption reveals some interesting patterns. In the binary case, 52 percent of the persistence is causal (true state dependence). The remaining sources for the time dependence in marijuana consumption are: i) persistence in time-varying utility shocks (18 percent); ii) persistent, observed individual characteristics, such as race, gender and family background (16 percent); and lastly, iii) persistent, unobserved heterogeneity (14 percent).

The estimated persistence probabilities for the ordered model suggest that time-invariant, unobserved heterogeneity plays a larger role in the persistence of intense marijuana consumption (40 percent of overall persistence is due to unobserved heterogeneity) and less so for moderate use (32 percent). Persistence in time-varying utility shocks and persistence due to time-invariant observed individual characteristics play a similar role to that obtained in the binary mode. Moreover, true or causal state dependence accounts for 47 percent of total persistence for moderate use, while it is less important for heavy consumption levels (33 percent). Most of the overall persistence in moderate consumption is due to structural state dependence (this result also applies when we consider consumption as a binary outcome), while for heavy consumption, most of the persistence is due to individual heterogeneity.

The rest of the paper is organized as follows. In the next section, we describe the data, and section 3 presents the econometric model and its results when considering marijuana consumption as a binary outcome. Section 4 is structured similarly but for the generalized model with ordered outcomes. Finally, section 5 concludes the paper with a brief summary.

3.2 Data

In this paper, we utilize data from the 1997 cohort of the National Longitudinal Survey of Youth (NLSY97), a nationally representative sample of five cohorts of males and females born between 1980 and 1984. The initial interview took place in 1997, and follow-up interviews were conducted annually until 2011, after which it became a biannual survey. NLSY97 gathers information in an event history format, in which dates are collected for the beginning and end of significant life events. In addition, there is detailed information on family background and income as well as on individual scholastic ability.

In our analysis, we remove individuals who were not part of the representative cross-sectional sample in 1997 (this removes oversamples of Blacks and Hispanics). In order to reduce potential initial conditions concerns, we also exclude all respondents who were born before 1983. Most of those born in 1983 were 13 years old at the time of the first survey, while most of those born in 1984 were 12 years old at that interview. We are then left with 1,589 individuals. Of these, 55 reported having used marijuana before the age of 13, and these were removed to avoid left censoring.

We also excluded individuals who did not provide valid information on the following: family income (at any point between 1997 and 2001), mother's age at birth, the family situation at the time of the survey (divorced parents or not), area of residence, number of siblings, mother's education and Armed Forces Qualification Test (AFQT) scores.[3] We exclude those with missing information on any of these variables since they are included as covariates in all model specifications.[4]

Finally, we remove respondents who did not provide any answers to questions related to marijuana use and those whom we only observed once. After these selections, the sample consists of 1,204 individuals.

We use the information on family income for each individual at ages 16 and 17, if available, and construct an average income measure. If income is only available for one of the years, the average income is replaced by that income. If no income information is available for these ages, we consider the income at earlier ages if available in order to minimize the number of individuals dropped because of missing income. We express income in the year 2000 dollars using the CPI for all urban consumers.

To derive measures of marijuana use, we compile information from questions like: 1) have you ever used marijuana?; 2) when did you start using marijuana?; 3) did you use marijuana during the year before the interview? And 4) On how many days have you used marijuana in the last 30 days? From the responses to these questions, we create individual annual indicators of marijuana use (and non-use) as well as indicators for intensity of use, conditional on use (less than ten days last month versus ten days or more). Responses to the first three questions are used to validate consistency in responses, while our outcome variables are derived from answers to the fourth question.

---

[3] AFQT scores consists of four components of the Armed Forces Vocational Aptitude Battery (ASVAB): Arithmetic Reasoning (AR), Mathematics Knowledge (MK), Word Knowledge (WK), and Paragraph Comprehension (PC). These scores have been used extensively in research on education using NLSY data. In this paper, we follow Belzil and Hansen (2020) and regress the scores on age and education, in order to adjust for age and educational differences at the time of the test, and use the standardized residual from that regression as the measure of cognitive ability.

[4] These variables are commonly included in empirical analysis of substance use. We decided not to include father's education in the list mainly because of the large number of missing values for this variable and the skewness in responses to questions about this across the sample (there is a higher fraction of missing among non-white respondents).

Table 3.1 presents the proportions of the sample that used marijuana at a given age. At age 13, 3.7 percent of the respondents used marijuana at least once. Three years later, at age 16, this had increased almost fivefold to 18.3 percent. After 16, the proportion of users increases until age 18, when it peaks and then declines to around 16 percent when respondents are in their 20s.

Table 3.1: Proportion of respondents using marijuana, by age

Age	Used marijuana	Numer of individuals
13	0.037	1,204
14	0.091	1,204
15	0.154	1,176
16	0.183	1,142
17	0.204	1,103
18	0.218	1,064
19	0.196	1,024
20	0.2	977
21	0.18	937
22	0.186	916
23	0.16	883
24	0.161	859
25	0.165	843
26	0.162	832

The entries in Table 3.1 do not reveal how respondents move in and out of marijuana use. In order to infer the degree of time persistence and the transitory nature of marijuana use, we show average (across individuals and time periods) conditional probabilities in Table 3.2. The entries show row percentages of the probability of using marijuana in year t, conditional on marijuana use in year $t-1$. The top row entries show that 91.5 percent of those who did not use marijuana in year $t-1$ continued to be non-users in year t, while 8.5 percent started using marijuana. Similarly, among those who used marijuana in year $t-1$, 63 percent continued using it in year t, while 37 percent stopped.[5]

[5]Deza (2015) reports similar proportions (an entry probability of 9.2 percent and a persistence probability of 67 percent (Table 3.2, panel B)) using NLSY97, despite different sample selections. She limited her sample to respondents with a valid state of residence at each wave between 1997 and 2007, i.e. a balanced panel. She also included the oversample of minorities available in NLSY97.

Table 3.2: Transition matrix

	Used marijuana in year t	
Used marijuana in year $t-1$	Yes	No
Yes	0.630	0.370
No	0.085	0.915

While the entries in Table 3.1 show how usage varies with age, the entries in Table 3.2 show the anatomy of usage in any year. That is how many start using it and how many stop. The focus of this paper is to analyze the persistence over time in marijuana use and estimate to what extent it is causal (or due to addiction) as opposed to persistence in observed and unobserved characteristics.

Table 3.3 shows average characteristics separately for individuals who never used marijuana and those who used it at least once over the sample period. Overall, males and Hispanics are somewhat overrepresented among users. In addition, at the interview date, the proportion of living with both biological parents is higher among the never-users (0.66) than among the users (0.57). There are no major differences in sample means between the two groups for other background variables - family income, mother's education, AFQT scores, mother's age at birth, urban residence, and the number of siblings. Lastly, half of our sample have used marijuana at least once. This is somewhat lower than the 57-58 percent reported in Deza (2015).

Table 3.3: Sample means, by marijuana use

	Never use	Used at least once
Male	0.49	0.55
Black	0.14	0.13
Hispanic	0.11	0.13
Intact family	0.66	0.57
Family income	\$66,191	\$65,429
Mother - high school graduate	0.33	0.34
Mother - attend college	0.53	0.53
AFQT	170.9	172
Mother's age at birth	26.4	26.2
Urban	0.71	0.75
Number of siblings	2.5	2.4
Peers	0.08	−0.08
Number of individuals	598	606

- Family income is expressed in year 2000 dollars.

Similar to earlier studies on substance use that utilize retrospective information, our measures of marijuana are subject to potential measurement error problems, specifically recall errors. However, unlike most of them (see, for instance, Van Ours and Williams (2009)), whose sample consists of respondents aged 25-50), the respondents in our sample were first asked about their marijuana use at a young age (age 12 or 13). We, therefore, believe the issue of recall errors is less serious in this paper than in many of the previous studies on this topic.

3.3 Binary Outcome

3.3.1 Estimation

In this paper we explore the persistence in marijuana use and its sources. Exploiting the longitudinal nature of the NLSY97 data, we analyze the dynamics of marijuana use (and non-use). Our empirical models are inspired by Heckman (1981) who derived a general framework for the analysis of discrete choices in discrete time. He showed that observed choices can be derived from latent variables, which in turn can be thought of as describing utility differences across alternatives. Hence, observed choices are outcomes of utility maximization. We follow Lee (1997) and Liu et al. (2012) who offers a description and assessment of generalized versions of Heckman's original framework.

Specifically, let y^*_{it} denote latent, unobserved utility differences, for individual i in period t, between using and not using marijuana

$$y^*_{i,t} = \Psi_{i,t} + \gamma y_{i,t-1} + \sigma\mu_i + \varepsilon_{i,t} \tag{3.1}$$

for $i = 1, ..., n; t = 1, ..., T_i$ and where $\Psi_{i,t} = X_i\beta + \kappa_1 (t - t_0) + \kappa_2 (t - t_0)^2$. If the utility difference is positive, individual i consumes marijuana in period t and the observed outcome is

$$y_{i,t} = \begin{cases} 1 \; if \; y^*_{i,t} > 0, \\ 0 \; if \; y^*_{i,t} \leq 0. \end{cases} \tag{3.2}$$

In our case, $y_{i,0} = 0$ as we start observing and modeling marijuana use at age 13. We include a fairly rich set of observable characteristics in X and assume that the error terms (μ_i) and $(\varepsilon_{i,t})$ are independent of X and across individuals. While μ_i is fixed over time, $\varepsilon_{i,t}$ is time-varying and possibly correlated over time. There are four possible sources of time persistence in marijuana use in equation 3.1: i) time-invariant observed characteristics (X_i); ii) true state dependence $(\gamma > 0)$; iii) time-invariant unobserved characteristics (μ_i); and iv) persistence in time-varying shocks $(\varepsilon_{i,t})$.

In equation 3.1, it is assumed that the dynamics of marijuana use can be fully captured by lagged choices $(y_{i,t-1})$. Alternatively, we can imagine that there is some memory in the process and that usage in previous periods may also have a direct or causal impact on current use. To allow for this, we consider a more general dynamic representation, described as the Polya model in Lee (1997), where the latent variable

$y_{i,t}^*$ is expressed as

$$y_{i,t}^* = \Psi_{i,t} + \gamma \sum_{j=1}^{t} \delta^{j-1} y_{i,t-j} + \sigma\mu_i + \varepsilon_{i,t} \tag{3.3}$$

for $i = 1, ..., n; t = 1, ..., T_i$ and where $\delta,\ [0,1]$ can be thought of as a discount factor. When $\delta = 0$, past choices beyond $t-1$ do not matter for the utility in period t whereas when $\delta = 1$, the impact of past choices do not fade with time.

We assume that $\varepsilon_{i,t} = \rho\varepsilon_{i,t-1} + \nu_{i,t}$, where ν_{it} are $i.i.d\ N(0,1)$, and consequently the choice probabilities involve multiple integrals. Following Lee (1997), we adopt the Geweke-Hajivassiliou-Keane (GHK) simulator and estimate the parameters in equations (1) and (2) using Maximum Simulated Likelihood. The joint probability for observed choices $y_{i,1}, ..., y_{i,T}$, conditional on X_i and μ_i is

$$\begin{aligned} Pr\left(y_{i,1}, ., y_{i,T} | X_i, \mu_i\right) = \int_{L_1}^{U_1} . \int_{L_T}^{U_T} & f\left(\varepsilon_{i,T} | \varepsilon_{i,T-1}, ., \varepsilon_{i,1}\right) \\ & f\left(\varepsilon_{i,T-1} | \varepsilon_{i,T-2}, ., \varepsilon_{i,1}\right) ... f\left(\varepsilon_{i,1}\right) d\varepsilon_T . d\varepsilon_1 \end{aligned} \tag{3.4}$$

where $f\left(\varepsilon_{i,t} | \varepsilon_{i,t-1}, .., \varepsilon_{i,1}\right)$ is the density of $\varepsilon_{i,t}$ conditional on past realizations of ε and the integral limits are

$$L_t = \begin{cases} -\left(\Psi_{i,t} + \gamma \sum_{j=1}^{t} \delta^{j-1} y_{i,t-j} + \sigma\mu_i\right) & if\ y_{i,t} = 1, \\ -\infty & if\ y_{i,t} = 0, \end{cases} \tag{3.5}$$

and

$$U_t = \begin{cases} \infty & if\ y_{i,t} = 1, \\ -\left(\Psi_{i,t} + \gamma \sum_{j=1}^{t} \delta^{j-1} y_{i,t-j} + \sigma\mu_i\right) & if\ y_{i,t} = 0, \end{cases} \tag{3.6}$$

Lee (1997) shows how the joint probability in equation 3.4 can be expressed using standard normal density and distribution functions and simulated using the GHK simulator. The sample likelihood then becomes

$$\mathcal{L} = \sum_{i=1}^{n} ln \left\{ \frac{1}{m} \sum_{j=1}^{m} \prod_{t=1}^{T_i} \Phi \left(D_{i,t} \left(\Psi_{i,t} + \gamma \sum_{j=1}^{t} \delta^{j-1} y_{i,t-j} + \sigma \mu_i^j + \rho \varepsilon_{i,t-1}^j \right) \right) \right\}, \quad (3.7)$$

where $D_{i,t} = 2y_{i,t} - 1$. The random disturbances $\varepsilon_{i,t}$ are recursively generated as described in Lee (1997).[6] The $\mu' s$ are generated from $N(0,1)$ random draws while the $\varepsilon' s$ are generated from functions of $U[0,1]$ draws. Lee (1997) provides Monte Carlo results for this and other dynamic specifications and concludes that this estimator generally performs well. Since we use an unbalanced panel, T_i varies between 2 and 14. We set $m = 100$.

3.3.2 Empirical Results

In this section, we present both parameter estimates and average partial effects of selected variables. We use a parametric bootstrap to estimate the standard errors of the average partial effects. Specifically, for each model we draw 100 vectors of parameter values from the estimated variance-covariance matrix. For each vector and variable of interest, we calculate a partial effect. The reported effects below are the average effects across the 100 draws and the standard errors of the effects are estimated using the standard deviation of the simulated effects.

3.3.2.1 Estimates and Average Partial Effects

Estimates from three alternative Probit specifications are presented in Table 3.4. This will allow us to analyze how the average partial effects depend on stochastic assumptions and specifications of the dynamic relationship of marijuana consumption.

[6] We provide a description of the generation of truncated random draws needed for the likelihood function in the Appendix D

Table 3.4: Selected estimates from binary probits.

	Model 1	Model 2	Model 3
Used marijuana in (t-1)	1.691	0.976	0.732
	(0.032)	(0.045)	(0.053)
σ	–	0.851	0.414
		(0.040)	(0.032)
ρ	–	–	0.220
			(0.083)
Male	0.113	0.157	0.133
	(0.028)	(0.063)	(0.050)
Intact family	-0.113	-0.244	-0.148
	(0.024)	(0.068)	(0.058)
AFQT	0.047	0.074	0.062
	(0.015)	(0.035)	(0.033)
Peers	-0.071	-0.115	-0.076
	(0.011)	(0.034)	(0.025)
AIC	$9,422$	$8,887$	$8,771$
LogL	$-4,695$	$-4,427$	$-4,368$

- Standard errors in parentheses.
- AIC is the Akaike Information Criteria.
- The dynamics of marijuana use in Models 1 and 2 are assumed to follow a first-order Markov structure.
-In Model 3, the dynamics is generalized to incorporate use prior to last year.
-Models 2 and 3 were estimated using simulated Maximum Likelihood with 100 simulation draws.

The entries in column one refer to a specification where dynamics in marijuana use is represented by a first-order Markov but with no time-invariant unobserved heterogeneity and no persistence in the time-varying shocks. In column two, we retain the assumption of a first-order Markov but allow for both unobserved heterogeneity and serial correlation in the time-varying shocks. Finally, in column three we generalize dynamics of marijuana use by incorporating marijuana use from periods before last year (see equation 3.3). We set δ to 0.7.

There is evidence of significant time dependence in marijuana use. The estimate in column one for marijuana use in the previous period (γ) is 1.691 and it is statistically significant. However, as discussed above, in this simplified model, all persistence in marijuana is captured by this parameter and it is therefore unlikely to represent the true (or causal) effect of past use on current use. Maintaining the same dynamic structure but allowing for another source of persistence has a dramatic (and expected) effect. The estimate in column two is 0.976, suggesting that the causal effect of past usage is seriously exaggerated in the naive specification in column one. Instead, a significant part of the observed persistence is due to time-invariant, unobserved heterogeneity with $\hat{\sigma}$ equal to 0.851.

The corresponding estimates reported in column three suggest important roles for all three sources of time dependence. The estimate of previous use (γ) is further reduced to 0.732 while $\hat{\sigma}$ equals 0.414. Further, $\hat{\rho}$ is significant and equals 0.220. At the bottom of Table 3.4, we report the Akaike Information Criteria (AIC) for each model specification and these favor the most general model presented in column three.

Regarding observable characteristics, the entries in Table 3.4 suggest that gender, family stability and size, cognitive skills and peer effects matter for marijuana use. The estimates associated with these variables are significant and generally similar across all three specifications while the estimates of the other included variables (shown in Table D.1) are not.

In Table 3.5 we show the average partial effects for selected variables. The average partial effects are estimated as $\hat{Pr}\left(y_{i,t}=1|y_{i,t-1}=1\right)-\hat{Pr}\left(y_{i,t}=1|y_{i,t-1}=0\right)$, and they are averaged across individuals and time periods. The first row shows the predicted difference in the probability of using marijuana between users and non-users in the previous period. According to these estimated effects - for the restrictive model with a first-order Markov dynamics, no unobserved heterogeneity and no serial persistence in the error terms - the probability of marijuana use in any given year is 47 percentage points higher if the person used marijuana the year before. This is a very large effect considering that the proportion of the sample that use marijuana at any given age very between 15 and 20 percent (after age 14, see Table 1). However, as we generalize the models, this conditional probability is reduced. In column two, the difference is 19.2 percentage points while in column three it has been reduced to 12.9 percentage points.[7]

Table 3.5: Average partial effects from binary probits.

	Model 1	Model 2	Model 3
Used marijuana in $(t-1)$	0.473	0.192	0.129
	(0.009)	(0.013)	(0.014)
Male	0.019	0.021	0.017
	(0.005)	(0.008)	(0.006)
Intact family	−0.018	−0.032	−0.018
	(0.004)	(0.009)	(0.008)
AFQT	0.008	0.010	0.007
	(0.002)	(0.005)	(0.004)
Peers	−0.012	−0.016	−0.010
	(0.002)	(0.004)	(0.003)

- Standard errors in parentheses.
- AIC is the Akaike Information Criteria.
- The dynamics of marijuana use in Models 1 and 2 are assumed to follow a first-order Markov structure.
- In Model 3, the dynamics is generalized to incorporate use prior to last year.
- Models 2 and 3 were estimated using simulated Maximum Likelihood with 100 simulation draws.
- A parametric bootstrap with 100 draws was used to estimate the standard errors.

The remaining entries in Table 3.5 show estimated marginal effects of the variables whose parameter estimates are statistically significant. Overall, and unlike the effect of past use, the magnitudes are similar across the different model specifications. For instance, the predicted probability of using marijuana is around two percentage points higher for males than for females while it is around two percentage points lower for students living with both biological parents at the time of the interview. Students with higher cognitive test scores (AFQT) have higher predicted probabilities of marijuana

[7]The average partial effect for Model 2 is a bit lower than the corresponding effect (25.1 percentage points) reported in Deza (2015). Her model, like the one in Model 2, ignores serial persistence in utility shocks and assume that a first-order Markov structure accurately captures dynamics in marijuana consumption.

use although the differences are small (a one standard deviation increase in test scores raise the probability with less than one percentage point). Finally, the effect of peers is just over one percentage point across all specifications suggesting that students with favorable peers are less likely to use marijuana.

3.3.2.2 Model Fit

We assess the model's ability to generate outcomes that match those observed in the data by predicting transition probabilities. In Table 3.6, we show the predicted transition matrix for marijuana use obtained by simulating outcomes generated by the estimates from the general Polya model (Model 3 in Table 3.4). The predicted conditional probabilities, which are averaged over individuals and time, match those in the data (presented in Table 3.2) well. For example, the probability of using marijuana in year t, conditional on using marijuana in year $t-1$, is 0.63 in the data and the predicted probability is 0.66. Moreover, the probability of using marijuana in year t, conditional on not using marijuana in year $t-1$ is 0.085 in the data while the predicted probability is 0.099.

Table 3.6: Predicted transition matrices

	Used marijuana in year t	
Used marijuana in $(t-1)$	Yes	No
Yes	0.661	0.339
	(0.159)	(0.159)
No	0.099	0.901
	(0.054)	(0.054)

- Average transition probabilities from simulation of outcomes using estimates from model 3 in Table 3.4 (the Polya model).
- Standard errors in parentheses.
- A parametric bootstrap with 100 draws was used to estimate the standard errors.

3.3.2.3 Sources of Persistence

In Table 3.7 we explore the anatomy of persistent marijuana use. The entries are obtained using estimates from the Polya model and in the first row, we replicate the the probability of using marijuana in year t, conditional on using marijuana in year $t-1$, from Table 3.6. This is the predicted persistence. In the second row, we remove the role of time-invariant unobserved heterogeneity by setting $\sigma = 0$ and the predicted probability drops from 0.661 to 0.567. Thus, removing time-invariant unobserved heterogeneity reduce persistence with 14 percent. In row three, we remove persistence in the time-varying utility shocks by setting $\rho = 0$ (in addition to setting $\sigma = 0$). The predicted persistence further drops to 0.449 indicating that this source of persistence contributes about 20 percent to the overall persistence.

Finally, in the last row, we also remove the effect of time-invariant observed characteristics and the time trend by setting $\beta = \kappa_1 = \kappa_2 = 0$ (in addition to fixing $\sigma = \rho = 0$). This further reduce the persistence from 0.449 to 0.345. The remaining persistence (52 percent of the total) is due to a causal or addictive effect of using marijuana in the previous period. Thus, a majority of the observed state dependence in marijuana consumption is causal although a large portion is due to persistence in utility shocks and heterogeneity. A similar finding is reported in Deza (2015).

Table 3.7: Sources of persistence

	Polya
(1) Predicted persistence	0.661
(2) Removing time-invariant unobserved heterogeneity	0.567
Proportion of total persistence - (2)/(1)	0.857
(3) Removing time-varying unobserved characteristics and (2)	0.449
Proportion of total persistence - (3)/(1)	0.679
(4) Removing observed characteristics, time trend and (3)	0.345
Proportion of total persistence - (4)/(1)	0.522

- The entries are derived using estimates from Model 3 in Table 3.4 and show $Pr\left(y_{i,t}=1|y_{i,t-1}=1\right)$.
In (2), we set $\sigma_u = 0$.
In (3), we set $\sigma_u = 0$; $\rho = 0$.
In (4), we set $\sigma_u = 0; \rho = 0; \beta = 0; \kappa_1 = 0; \kappa_2 = 0$.

3.4 Ordered Outcome

The results so far are based on the dichotomy of marijuana use with no separation between occasional or moderate consumption and more intense, regular use. This is arguably restrictive and to allow for different effects depending on the intensity of consumption, we generalize the model described above to include multiple, ordered outcomes.[8]

3.4.1 Estimation

Specifically, let $c^*_{i,t}$ denote latent, unobserved utility of marijuana consumption for individual i in period t

$$c^*_{i,t} = \Psi_{i,t} + \gamma_1 \sum_{j=1}^{t} \delta^{j-1} 1\left(c_{i,t-1} = 1\right) + \gamma_2 \sum_{j=1}^{t} \delta^{j-1} 1\left(c_{i,t-1} = 2\right) + \sigma\mu_i + \varepsilon_{i,t}, \tag{3.8}$$

[8] Honoré et al. (2021) derive a generalized method of moments estimator for a dynamic ordered Logit model with fixed effects, assuming time independence of the utility shocks. We argue that since we observe the initial conditions, the argument for using a fixed effects estimator instead of a random effects estimator (like we do) is weaker.

for $i = 1, ..., nandt = 1, ..., T_i$, where $\mathit{\Psi}_{i,t} = X_i\beta + \kappa_1 (t - t_0) + \kappa_2 (t - t_0)^2$. $1(.)$ is an indicator function that equals one if the argument is true and zero otherwise. If utility is below a certain level (θ_1), the individual is not consuming marijuana in period t. If utility exceeds (θ_1) but is below (θ_2), the individual consumes a moderate amount of marijuana in period t and finally, if utility exceeds (θ_2), the individual is a heavy user. Thus, the observed outcome $(c_{i,t})$ is

$$c_{i,t} = \begin{cases} 0 \; if \; c^*_{i,t} \leq \theta_1, \\ 1 \; if \; \theta_1 < c^*_{i,t} \leq \theta_2, \\ 2 \; if \; c^*_{i,t} > \theta_2. \end{cases} \tag{3.9}$$

As mentioned above in the binary case, $c_{i,0} = 0$ since we start observing and modeling marijuana use at age 13. We maintain the assumptions that the error terms (μ_i) and (ε_{it}) are independent of X and across individuals, μ_i is $i.i.d.$ $N(0, 1)$ and fixed over time while $\varepsilon_{i,t} = \rho\varepsilon_{i,t-1} + \nu_{i,t}$, where $\nu_{i,t}$ are $i.i.d$ $N(0, 1)$. We define $c_{i,t} = 0$ if the person did not use marijuana in period t, $c_{i,t} = 1$ if the person used marijuana less than 10 times per month in period t (moderate use) and $c_{i,t} = 2$ if the person used marijuana 10 times or more per month in period t (heavy use).

Given the stochastic assumptions and the assignment rule above, the probabilities of observed outcomes are then

$$\begin{aligned} Pr(c_{i,t} = 0|c_{i,t-1}) &= \Phi(\theta_1 - \lambda_{i,t}) = \mathit{\Lambda}_0, \\ Pr(c_{i,t} = 1|c_{i,t-1}) &= \Phi(\theta_2 - \lambda_{i,t}) - \Phi(\theta_1 - \lambda_{i,t}) = \mathit{\Lambda}_1, \\ Pr(c_{i,t} = 2|c_{i,t-1}) &= 1 - \Phi(\theta_2 - \lambda_{i,t}) = \mathit{\Lambda}_2, \end{aligned} \tag{3.10}$$

where

$$\lambda_{i,t} = \mathit{\Psi}_{i,t} + \gamma_1 \sum_{j=1}^{t} \delta^{j-1} 1(c_{i,t-1} = 1) + \gamma_2 \sum_{j=1}^{t} \delta^{j-1} 1(c_{i,t-1} = 2) + \sigma\mu_i + \rho\varepsilon_{i,t-1}. \tag{3.11}$$

We again adopt the Geweke-Hajivassiliou-Keane (GHK) simulator and estimate the parameters in equation 3.8 using Maximum Simulated Likelihood. The sample likelihood is an adjusted version of the one presented in equation 3.7 above

$$\mathcal{L} = \sum_{i=1}^{n} ln \left\{ \frac{1}{m} \sum_{j=1}^{m} \prod_{t=1}^{T_i} \Lambda_0{}^{I(c_{it}=0)} \Lambda_1^{I(c_{it}=1)} \Lambda_2^{I(c_{it}=2)} \right\}. \tag{3.12}$$

The random disturbances $\varepsilon_{i,t}$ are generated recursively, similar to the binary case, and the $\mu' s$ are generated from $N(0,1)$ random draws while the $\varepsilon' s$ are generated from functions of $U[0,1]$ draws.[9] We set $m = 100$.

3.4.2 Empirical Results

3.4.2.1 Descriptive Statistics

The proportions of the sample that used marijuana at a given age, by intensity level, are presented in Table 3.8. At age 13, of the 3.7 percent of the respondents who used marijuana at least once, a majority (73 percent) used it occasionally (less than 10 days during the 30 days preceding the survey date). Three years later, at age 16, the proportion of intense users, among all users, increase to 33 percent. In fact, the proportion of intense users, among all users, increase with age and reach over 60 percent at age 26. This suggests a higher degree of persistence among the intense users.

[9] See the Appendix D for details.

Table 3.8: Proportion of respondents using marijuana, by age

Age	Did not use marijuana	Used marijuana less than 10 days	Used marijuana 10 days or more	Number of individuals
13	0.963	0.027	0.010	1,204
14	0.909	0.073	0.018	1,204
15	0.846	0.107	0.047	1,176
16	0.817	0.122	0.061	1,142
17	0.796	0.119	0.085	1,103
18	0.782	0.116	0.102	1,064
19	0.804	0.104	0.093	1,024
20	0.800	0.107	0.092	977
21	0.820	0.099	0.081	937
22	0.814	0.094	0.092	916
23	0.840	0.079	0.080	883
24	0.839	0.079	0.081	859
25	0.835	0.077	0.088	843
26	0.838	0.064	0.099	832

The entries in Table 3.9show the degree of time persistence and the transitory nature of marijuana use, conditional on intensity of consumption. Like before, we show average (across individuals and time periods) conditional probabilities and the entries show row percentages of the probability of consuming a certain level of marijuana in year t, conditional on marijuana use in year $t-1$. The top row entries show, like before, that 91.5 percent of those who did not use marijuana in year $t-1$ continued to be non-users in year t. Among the remaining non-users, 6.4 percent started consuming marijuana at a moderate intensity level while 2.1 percent (a quarter of those who started using marijuana) used marijuana intensively (used it at least 10 days or more during the 30 days preceding the survey date). Among those who used marijuana moderately in year $t-1$, almost half stopped consuming it in year t while 16 percent increased their consumption the following year. Only 34 percent continued with moderate use, suggesting a transitory nature among occasional or moderate users. The entries in the last row show that 20 percent of the intense users in period $t-1$ stopped using marijuana in period t while 16.6 percent reduced their consumption (but kept consuming). However, the majority (63.5 percent) continued their intense level of consumption the following year (in year t).

Table 3.9: Transition matrix

	days of marijuana use last month in year t		
	0	1-9	10 or more
days of marijuana use last month in year $t-1$			
0	0.915	0.064	0.021
1-9	0.497	0.339	0.164
10 or more	0.198	0.166	0.635

- Row percentages

3.4.2.2 Estimates and Average Partial Effects

Estimates from the ordered Probit Polya model (the likelihood presented in equation 3.12) are shown in Table 3.10. Similar to the binary case, we set δ to 0.7. The model includes the same set of observed characteristics as the ones for the binary case but we report only a subset of the associated estimates in Table 3.10 (those that are statistically significant). The full set of estimates are provided in Table D.2 in Appendix D.

The estimates in the first two rows suggest existence of true or causal time dependence in outcomes and this dependence is stronger for intense marijuana use. The estimates are 0.432 and 0.786 for moderate and heavy use, respectively. We will illustrate how these estimates translate into average partial effects and predicted transition probabilities below. The estimates for male, intact family and peers are similar in magnitude (and statistical significance) to those obtained in the binary case (see column 3 of Table 3.4). The standard deviation of the persistent unobserved heterogeneity term, $\hat{\sigma}$, is 0.569, again similar to the estimate in the binary model. Finally, there is evidence of serial persistence in the error terms (ε_{it}) as $\hat{\rho}$ is significant and equals 0.300.

Table 3.10: Selected estimates from an ordered probit polya model

	Estimate	Standard error
Marijuana, 1-9 days $(t-1)$	0.432	0.047
Marijuana, 10+ days $(t-1)$	0.786	0.051
Male	0.143	0.047
Intact family	-0.183	0.054
AFQT	0.056	0.031
Peers	-0.088	0.026
σ	0.569	0.06
ρ	0.3	0.025
θ_1	1.735	0.182
θ_2	2.556	0.186
	-5623	

- The specification included additional observed characteristics (the same list as in Table 3.4).
- The remaining parameter estimates and standard errors are presented in Table D.2 in Appendix D.

In Table 3.11 we show the average partial effects for selected variables. The first two rows show the predicted difference in the probability of using marijuana at a moderate level when we condition on different consumption levels in the previous time period. The first effect is the difference in conditional probabilities of moderate consumption when we condition on moderate versus no consumption in the previous time period while the second effect conditions on moderate and heavy use instead. The former effect (moderate versus no consumption) is 0.046 while the second effect is -0.051. That is, the probability of consuming moderate levels of marijuana in year t is 4.6 percentage points higher if the person consumed the same level of marijuana in year $t-1$, relative to not using any marijuana in year $t-1$. While the magnitude of this effect is smaller than the one obtained in the binary case, it constitutes a relative effect that is close to 50 percent, given the observed moderate consumption rates observed in the data. The negative effect for moderate versus heavy usage suggests a higher probability of moderate use in year t for those with a heavy consumption in the previous year compared to those with moderate consumption.

Table 3.11: Average partial effects from selected variables on the probability of moderate and heavy marijuana consumption.

	Moderate	Heavy
$Pr\left(y^m_{i,t}=1\|y^m_{i,t-1}=1\right)-Pr\left(y^m_{i,t}=1\|y^n_{i,t-1}=1\right)$	0.046	-
	(0.001)	
$Pr\left(y^m_{i,t}=1\|y^m_{i,t-1}=1\right)-Pr\left(y^m_{i,t}=1\|y^h_{i,t-1}=1\right)$	-0.051	-
	(0.001)	
$Pr\left(y^h_{i,t}=1\|y^h_{i,t-1}=1\right)-Pr\left(y^h_{i,t}=1\|y^n_{i,t-1}=1\right)$	-	0.043
		(0.001)
$Pr\left(y^h_{i,t}=1\|y^h_{i,t-1}=1\right)-Pr\left(y^h_{i,t}=1\|y^m_{i,t-1}=1\right)$	-	0.027
		(0.001)
Male	0.013	0.004
	(0.001)	(0.001)
Intact family	-0.016	-0.005
	(0.001)	(0.001)
AFQT	0.003	0.001
	(0.001)	(0.001)
Peers	-0.005	-0.001
	(0.001)	(0.001)

- A parametric bootstrap with 100 draws was used to estimate the standard errors of the average partial effects.

In rows three and four we present the corresponding probability differences for heavy consumption levels. The first effect is the difference in conditional probabilities of heavy consumption when we condition on heavy versus no consumption in the previous time period while the second effect conditions on heavy and moderate use instead. The former effect equals 0.043 and is similar to the one estimated for moderate use. The second effect is smaller, 0.027. That is, the probability of consuming heavy levels of marijuana in year t is 4.3 percentage points higher if the person consumed the same level of marijuana in year $t-1$, relative to not using any marijuana in year $t-1$. Again, while the magnitude of this effect is smaller than the one obtained in the binary case, it constitutes a relative effect that is close to 50 percent, given the observed moderate consumption rates observed in the data.

The remaining entries in Table 3.11 show estimated marginal effects of the variables whose parameter estimates are statistically significant. For all four variables (male, intact family, afqt and peers), the average partial effects are larger in absolute value for moderate use than for heavy use. For example, the predicted probability of using a moderate level of marijuana is 1.3 percentage points higher for males than for females while it is only 0.4 percentage points higher in the heavy consumption case. A similar difference applies to the impact of living with both biological parents at the time of the interview. While youth in intact families are less likely to use any level of marijuana, the strength of the effect is weaker for heavy use (-0.005 versus -0.016 for moderate use).

3.4.2.3 Model Fit

Similar to the binary case presented above, we assess the model's ability to generate outcomes that match those observed in the data by predicting transition probabilities. In Table 3.12, we show the predicted transition matrix for marijuana use obtained by simulating outcomes generated by the estimates from the ordered Polya model. The predicted conditional probabilities, which are averaged over individuals and time, match those in the data (presented in Table 3.9) reasonably well. For example, the probability of not using marijuana in year t, conditional on not using marijuana in year $t-1$ is 0.915 in the data and the predicted probability is 0.92. The predicted entry probabilities, going from non-use to moderate or intense use, also match those in the data well. The second row entries show probabilities of various use conditional on moderate use in period $t-1$. The predicted exit (or stopping) probability is 0.551 compared to 0.497 in the data. However, the model underestimates the probability of remaining a moderate user somewhat (0.254 versus 0.339) and slightly exaggerates the transition from moderate to intense use (0.195 versus 0.164). Conditional on heavy use, the predicted probabilities are similar to those in the data, especially the probability of remaining an intense user (0.598 versus 0.635 in the data). Overall, the model generates predicted transition matrix entries that match those in the data well.

Table 3.12: Model fit: Transition matrix

	Days of marijuana use last month in year t		
	0	1-9	10 or more
Days of marijuana use last month in year $t-1$			
0	0.923	0.06	0.017
	(0.003)	(0.002)	(0.001)
1-9	0.551	0.254	0.195
	(0.002)	(0.003)	(0.003)
10 or more	0.183	0.219	0.598
	(0.004)	(0.003)	(0.006)

- Row percentages.

3.4.2.4 Sources of Persistence

In Table 3.13 we replicate the analysis on the anatomy of persistent marijuana use but generalize it to allow differential impacts on moderate and heavy use. The entries in column one refers to moderate use, $\hat{Pr}\left(y^m_{i,t} = 1 | y^m_{i,t-1} = 1\right)$, while those in column two refer to heavy use, $\hat{Pr}\left(y^h_{i,t} = 1 | y^h_{i,t-1} = 1\right)$. They are obtained using estimates from the ordered Polya model and in the first row, we replicate the the probabilities of marijuana consumption in year t, conditional on the same intensity level of marijuana consumption in year $t-1$, from Table 3.12. In the second row, we remove the role of time-invariant unobserved heterogeneity by setting $\sigma = 0$. The predicted probability drops marginally from 0.254 to 0.173 in the moderate case and from 0.598 to 0.355 in the heavy case. Thus, persistent unobserved heterogeneity contributes significantly to time dependence in both types of marijuana consumption, by 32 percent for moderate consumption levels and by just over 40 percent for heavy use.

Table 3.13: Sources of persistence

	Persistence	
	Moderate	Heavy
(1) Predicted persistence	0.254	0.598
(2) Removing time-invariant unobserved heterogeneity	0.173	0.355
Proportion of total persistence - (2)/(1)	0.68	0.594
(3) Removing time-varying unobserved characteristics and (2)	0.156	0.255
Proportion of total persistence - (3)/(1)	0.615	0.427
(4) Removing observed characteristics, time trend and (3)	0.118	0.196
Proportion of total persistence - (4)/(1)	0.467	0.328

- The entries are derived using estimates from the model presented in Table 3.10 and show $Pr\left(y^j_{i,t} = 1 | y^j_{i,t-1} = 1\right)$, j=Moderate,Heavy.
- A parametric bootstrap with 100 draws was used to estimate the standard errors. In (2), we set $\sigma_u = 0$ and in (3), we set $\sigma_u = 0$; $\rho = 0$.
In (4), we set $\sigma_u = 0$; $\rho = 0$; $\beta = 0$; $\kappa_1 = 0$; $\kappa_2 = 0$.

In row three, we remove persistence in the time-varying utility shocks by setting $\rho = 0$ (in addition to setting $\sigma = 0$). The predicted persistence further drops to 0.156 for the moderate case and to 0.255 for the intense case. This source of persistence contributes about 7 percent to the overall persistence for both moderate levels of marijuana use and 17 percent for heavy levels. Finally, in the last row, we also remove the effect of time-invariant observed characteristics and the time trend by setting $\beta = \kappa_1 = \kappa_2 = 0$. This further reduces the persistence from 0.156 to 0.118 in the moderate case and from 0.255 to 0.196 in the intense case. The remaining persistence (47 percent of the total for moderate use and 33 percent of the total for intense use) is due to a causal or addictive effect of using marijuana in the previous period.

That is, most of the overall persistence in moderate consumption is due to structural state dependence (this result also applies when we consider consumption as a binary outcome) while for heavy consumption, most of the persistence is due to individual heterogeneity.

3.5 Conclusions

In this paper we provide new evidence on the persistence of marijuana use among American youth. This topic is important for many reason, one being the fact that marijuana consumption among teenagers is inversely related to many successful future labor market outcomes. It is perhaps more important than ever given the recent legalization of recreational marijuana use in many jurisdictions. Moreover, according to 2018 results on monitoring the future from the National Institute on Drug Abuse, marijuana use were at historic highs in 2018, both among college and non-college peers.

The previous literature on persistence of marijuana consumption is limited. A notable exception is Deza (2015) who estimate a dynamic discrete choice model of alcohol, marijuana and hard drugs use and focus on the state dependence in these, as well as dependence across different drugs. While our paper share many features with Deza (2015), there are also important differences. Unlike her, we allow for persistence in the utility shocks, in addition to persistence generated from time-invariant unobserved heterogeneity and pure or causal state dependence. Further, we specify the dynamics in marijuana use in a more flexible way and do not limit it to the inclusion of a one-period lag. Perhaps most importantly, in the second part of the paper, we distinguish between different intensity levels of marijuana consumption. Instead of using a binary outcome (used or not), we code moderate use (consumption during 1-10 days last month) separately from heavy use (consumption during 10 days or more last month). We show that moderate consumption is transitory and less persistent than heavy use. A significant fraction in the data (16.4 percent) of moderate users transit to heavy use in the next period while an even larger share (49.7 percent) stop using marijuana next period.

The estimated average partial effects show that previous consumption significantly increase the probability of current consumption. We show that these effects exist for all consumption levels but are severely exaggerated in models that ignores persistence in utility shocks and restricts the form of dynamics. However, even in the most general model specifications, the partial effects suggest that the probability of consuming marijuana now increase by a factor of 1.5 when we change the status of previous consumption from none to moderate or heavy. This finding is robust towards aggregation of marijuana consumption.

We also disaggregate overall persistence into four components and show the relative contribution of each. The results show that persistent unobserved heterogeneity plays a large role in persistence of heavy marijuana consumption (40 percent of overall persistence is due to unobserved heterogeneity) and less so for moderate use (32 percent is due to unobserved heterogeneity). Persistence in time-varying random shocks also play a significant role and its importance is similar that observed for persistent observed individual characteristics. Finally, true or causal state dependence is important for both intensity levels, but more so for moderate consumption (47 and 33 percent, respectively).

The results for moderate use are similar to those obtained in the binary case where there is no distinction between occasional and intense consumption. These results are also similar to those found in Deza (2015). However, by ignoring the possibility that structural persistence is a function of the level of consumption, the role of causal state dependence may be exaggerated. This in turn may lead to misguided policy recommendations as the risk of addictive behavior may be overstated.

We believe the framework and results provided in this paper will serve as a catalyst for further work in this important area of economics and health. For example, we have restricted the state dependence to be constant across individuals. It would be interesting to investigate if there are differences in persistence between males and females as well as across racial groups. Moreover, we have not considered the consumption of alcohol and cigarettes in this paper but this could be an interesting avenue for future research, building also on the work of Deza (2015). In a companion paper where we estimate the effect of marijuana use on educational attainment, Hansen and Davalloo (2022), we find that age of marijuana initiation is and important determinant of the effect. It may also impact the persistence of marijuana use. These are all topics for extensions of this paper that we plan to pursue in the near future.

References

Abbring, J. H. and Van Den Berg, G. J. (2003). The identifiability of the mixed proportional hazards competing risks model. *Journal of the Royal Statistical Society B*, 76(2):325–330.

Abbring, J. H., Van Den Berg, G. J., and Van Ours, J. C. (2005). The effect of unemployment insurance sanctions on the transition rate from unemployment to employment. *Economic Journal*, 115(505):602–630.

Aït-Sahalia, Y. and Xiu, D. (2019). A Hausman test for the presence of market microstructure noise in high frequency data. *Journal of Econometrics*, 211(1):176–205.

Arifin, W. N. (2017). Exploratory factor analysis and Cronbach ' s alpha. Technical report.

Baltagi, B. H., Bresson, G., and Etienne, J. M. (2015). Hedonic housing prices in Paris: an unbalanced spatial lag pseudo-panel model with nested random effects. *Journal of Applied Econometrics*, 30(3):509–528.

Bartlett, M. (1951). An inverse matrix adjustment arising in discriminant analysis. *The Annals of Mathematical Statistics*, 22:107–111.

Belzil, C. and Hansen, J. (2020). The evolution of the US family income-schooling relationship and educational selectivity. *Journal of Applied Econometrics*, (May):1–19.

Bernard, J. T., Bolduc, D., and Yameogo, N. D. (2011). A pseudo-panel data model of household electricity demand. *Resource and Energy Economics*, 33:315–325.

Beverly, H. K., Castro, Y., and Opara, I. (2019). Age of first marijuana use and its impact on education attainment and employment status. *Journal of Drug Issues*, 49(2):228–237.

Birkeland, S., Murphy-Graham, E., and Weiss, C. (2005). Good reasons for ignoring good evaluation: The case of the drug abuse resistance education (D.A.R.E.) program. *Evaluation and Program Planning*, 28(3):247–256.

Blundell, R., Browning, M., and Meghir, C. (1994). Consumer demand and the life-cycle allocation of household expenditures. *The Review of Economic Studies*, 61(1):57–80.

Bond, L., Patton, G., Glover, S., Carlin, J. B., Butler, H., Thomas, L., and Bowes, G. (2004). The Gatehouse Project: Can a multilevel school intervention affect emotional wellbeing and health risk behaviours? *Journal of Epidemiology and Community Health*, 58(12):997–1003.

Boruch, R., Crane, J., Ellwood, D., Gueron, J., Haskins, R., Kesslerm, D., Lee, J., Ravitch, D., Robinson, L., Sawhill, I., Seligman, M., Slavin, R., and Zill, N. (2003). Bringing evidence-driven progress to crime and substance-abuse policy: A recommended federal strategy. Technical report, The Council For Excellence In Government, Washington, DC.

Botvin, G. J., Baker, E., Dusenbury, L., Botvin, E. M., and Diaz, T. (1995). Long-term follow-up results of a randomized drug abuse prevention trial in a white middle-class population.

Botvin, G. J., Baker, E., Dusenbury, L., Tortu, S., and Botvin, E. M. (1990a). Preventing adolescent drug abuse through a multimodal cognitive-behavioral approach: results of a 3-year study. *Journal of consulting and clinical psychology*, 58(4):437–446.

Botvin, G. J. and Griffin, K. W. (2006). Drug abuse prevention curricula in schools.

In Sloboda, Z. and Bukoshi, W. J., editors, *Handbook of drug abuse prevention*, chapter 3, pages 45–74. Springer.

Botvin, J., Baker, E. L. I., Filazzola, A. D., and Botvin, E. M. (1990b). A cognitive-behavioural approach to substance abuse prevention: One-year follow-up. *Addictive Behaviors*, 15(Da 02835):47–63.

Bray, J. W., Zarkin, G. A., Ringwalt, C., and Qi, J. (2000). The relationship between marijuana initiation and dropping out of high school. *Health Economics*, 9:9–18.

Brewer, M. J., Butler, A., and Cooksley, S. L. (2016). The relative performance of AIC, AICC and BIC in the presence of unobserved heterogeneity. *Methods in Ecology and Evolution*, 7(6):679–692.

Brook, J. S., Stimmel, M. A., Zhang, C., and Brook, D. W. (2008). The association between earlier marijuana use and subsequent academic achievement and health problems: A longitudinal study. *American Journal on Addictions*.

Browning, M., Deaton, A., and Irish, M. (1985). A profitable approach to labor supply and commodity demands over the life-cycle. *Econometrica*, 53(3):503–544.

Carro, J. M. and Traferri, A. (2014). State dependence and heterogeneity in health using a bias-corrected fixed-effects estimator. *Source: Journal of Applied Econometrics*, 29(2):181–207.

Caulkins, J. P., Pacula, R. L., Paddock, S., and Chiesa, J. (2002). *School-based drug prevention: What kind of drug use does it prevent?*, volume 30. RAND.

Cerdá, M., Moffitt, T. E., Meier, M. H., Harrington, H. L., Houts, R., Ramrakha, S., Hogan, S., Poulton, R., and Caspi, A. (2016). Persistent cannabis dependence and alcohol dependence represent risks for midlife economic and social problems: A longitudinal cohort study. *Clinical Psychological Science*, 4(6):1028–1046.

Chan, G. C., Becker, D., Butterworth, P., Hines, L., Coffey, C., Hall, W., and Patton, G. (2021). Young-adult compared to adolescent onset of regular cannabis use: A

20-year prospective cohort study of later consequences. *Drug and Alcohol Review*, 40(4):627–636.

Chatterjee, S. and Hadi, A. S. (2012). *Regression analyses by example.*

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

Cobb-Clark, D. A., Kassenboehmer, S. C., Le, T., McVicar, D., and Zhang, R. (2013). "High"-school: The relationship between early marijuana use and educational outcomes.

Comeau, N., Stewart, S. H., and Loba, P. (2001). The relations of trait anxiety, anxiety sensitivity, and sensation seeking to adolescents' motivations for alcohol, cigarette, and marijuana use. *Addictive Behaviors.*

Costello, A. B. and Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Education*, 10:1–9.

Čurová, V., Orosová, O., Abrinková, L., and Štefaňáková, M. (2021). The effectiveness of drug use prevention programs on substance use among Slovak schoolchildren. In Pracana, C. and Wang, M., editors, *Psychology applications developments*, volume VII, chapter 12, pages 129–142. InSciencePress.

David Card, B. and Hyslop, D. R. (2005). Estimating the effects of a time-limited earnings subsidy for welfare-leavers. *Econometrica*, 73(6):1723–1770.

Deaton, A. (1985). Panel data from time series of cross-sections. *Journal of Econometrics*, 30:109–126.

Deza, M. (2015). Is there a stepping stone effect in drug use? Separating state dependence from unobserved heterogeneity within and between illicit drugs. *Journal of Econometrics*, 184(1):193–207.

Di Forti, M., Quattrone, D., Freeman, T., Tripoli, G., Gayer-Anderson, C., Quigley, H., and Rodriguez, V. (2019). The contribution of cannabis use to variation in the incidence of psychotic disorder across Europe (EU-GEI): A muticentre case-control study. *Lancet Psychiatry*, 6:427–436.

Dryfoos, J. (1990). *Adolescents at risk: Prevalence and prevention.* Oxford University Press, New York.

Dziak, J. J., Coffman, D. L., Lanza, S. T., and Runze, L. (2012). Sensitivity and specificity of information criteria. Technical Report 119, The Pennsylvania State University.

Ennet, S. T., Tobler, N. S., Ringwalt, C. L., and Flewelling, R. L. (1994). How effective is drug abuse resistance education? A meta-analysis of project DARE outcome evaluations. *American Journal of Public Health*, 84(9):1394–1401.

Ewing, S. W., Venner, K. L., Mead, H. K., and Bryan, A. D. (2011). Exploring racial/ethnic differences in substance use: A preliminary theory-based investigation with juvenile justice-involved youth. *BMC Pediatrics*, pages 1–10.

Ewing, S. W. F., Filbey, F. M., Loughran, T. A., Chassin, L., and Piquero, A. R. (2015). Which matters most ? Demographic, neuropsychological, personality, and situational factors in long-term marijuana and alcohol trajectories for justice-involved male youth. *Psychology of Addictive Behaviors*, 29(3):603–612.

Fergusson, D. M. and Boden, J. M. (2008). Cannabis use and later life outcomes. *Addiction*, 103:969–976.

Fergusson, D. M., Norwood, L. J., and Beautrais, A. L. (2003). Cannabis and educational achievement. *Addiction*, 98(12):1681–1692.

Flay, B. R. and Collins, L. M. (2005). Historical review of school-based randomized trials for evaluating problem behavior prevention programs. Technical Report July, American Academy of Political and Social Science.

Flay, B. R., Graumlich, S., Segawa, E., Burns, J. L., and Holliday, M. Y. (2004). Effects of 2 prevention programs on high-risk behaviors among African American youth. *Archives of Pediatrics and Adolescent Medicine*, 158(4):377–384.

Fletcher, A., Bonell, C., and Hargreaves, J. (2008). School effects on young people's drug use: A systematic review of intervention and observational studies. *Journal of Adolescent Health*, 42(3):209–220.

Gabrhelik, R., Duncan, A., Lee, M. H., Stastna, L., Furr-Holden, C. D. M., and Miovsky, M. (2012). Sex specific trajectories in cigarette smoking behaviors among students participating in the unplugged school-based randomized control trial for substance use prevention. *Addictive Behaviors*, 37(10):1145–1150.

Gardes, F., Duncan, G. J., Gaubert, P., Gurgand, M., and Starzec, C. (2005). Panel and pseudo-panel estimation of cross-sectional and time series elasticities of food consumption: The case of U.S. and Polish data. *Journal of Business and Economic Statistics*, 23(2):242–253.

Gerstein, D. R. and Green, L. W. (1993). *Preventing drug abuse: What do we know?* National Academies Press.

Gonzalez, R. and Swanson, J. M. (2012). Long-term effects of adolescent-onset and persistent use of cannabis.

Gorsuch, R. L. (1988). Exploratory factor analysis. In *Handbook of Multivariate Experimental Psychology*, pages 231–258. Plenum Press, New York.

Grant, J. D., Lynskey, M. T., Scherrer, J. F., Agrawal, A., Heath, A. C., and Bucholz, K. K. (2010). A cotwin-control analysis of drug use and abuse/dependence risk associated with early-onset cannabis use. *Addictive Behaviors*, 35:35–41.

Guillerm, M. (2017). Pseudo-panel methods and an example of application to household wealth data. *Economie et Statistique / Economics and Statistics*, 491-492:109–130.

Hall, W. and Degenhardt, L. (2009). Adverse health effects of non-medical cannabis use.

Ham, J. C. and Lalonde, R. J. (1996). The effect of sample selection and initial conditions in duration models: Evidence from experimental data on training. *Econometrica*, 64(1):175.

Hansen, J. and Davalloo, G. (2022). Heterogeneous effects of marijuana use on educational attainment.

Hansen, J. and Lofstrom, M. (2009). The dynamics of immigrant welfare and labor market behavior. *Source: Journal of Population Economics*, 22(4):941–970.

Hansen, W. B. (1992). School-based substance abuse prevention: A review of the state of the art in curriculum, 1980–1990. *Health Education Research*, 7(3):403–430.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 46(6):1251–1271.

Heckman, J. (1981). Heterogeneity and state dependence. Technical report, University of Chicago Press, Chicago.

Homel, J., Thompson, K., and Leadbeater, B. (2014). Trajectories of marijuana use in youth ages 15–25: Implications for postsecondary education experiences. *Journal of Studies on Alcohol and Drugs*, 75(4):674–683.

Honoré, B. E., Muris, C., and Weidner, M. (2021). Dynamic ordered panel logit models.

Hyslop, D. R. (1999). State dependence, serial correlation and heterogeneity in intertemporal labor force participation of married women. *Econometrica*, 67(6):1255–1294.

Inoue, A. (2008). Efficient estimation and inference in linear pseudo-panel data models. *Journal of Econometrics*, 142(1):449–466.

Jessor, R. and Jessor, S. L. (1977). *Problem behavior and psychosocial development: A longitudinal study of youth.* Academic Press, New York.

Jolliffe, I. (2011). *Principal component analysis.* In: Lovric M. (eds) International Encyclopedia of Statistical Science. Springer, Berlin, Heidelberg.

Kaiser, H. F. and Rice, J. (1974). Little jiffy, mark IV. *Educational and Psychological Measurement*, 34(1974):111–117.

Keane, M. P. and Wolpin, K. I. (1997). The career decisions of young men. *Journal of Political Economy*, 105(3):473–522.

Keyes, K. M., Rutherford, C., and Miech, R. (2019). Historical trends in the grade of onset and sequence of cigarette, alcohol, and marijuana use among adolescents from 1976–2016: Implications for "Gateway" patterns in adolescence. *Drug and Alcohol Dependence*, 194(June 2018):51–58.

Kuha, J. (2004). AIC and BIC: Comparisons of assumptions and performance. *Sociological Methods and Research*, 33(2):188–229.

Lee, L.-F. (1997). Simulated maximum likelihood estimation of dynamic discrete choice statistical models Some Monte Carlo results. *Journal of Econometrics*, 82:1–35.

Lev-Ran, S., Roerecke, M., Le Foll, B., George, T. P., Mckenzie, K., and Rehm, J. (2014). The association between cannabis use and depression: A systematic review and meta-analysis of longitudinal studies. *Psychological Medicine*, 44:797–810.

Liu, X., Kagel, J., and Lee, L.-F. (2012). Dynamic discrete choice models with lagged social interactions: With an application to a signaling game experiment. *Journal of Applied Econometrics*, 27:1037–1058.

McCaffrey, D. F., Pacula, R. L., Han, B., and Ellickson, P. (2010). Marijuana use and high school dropout: The influence of unobservables. *Health Economics*, 19(11):1281–1299.

McKenzie, D. J. (2004). Asymptotic theory for heterogeneous dynamic pseudo-panels. *Journal of Econometrics*, 120(2):235–262.

Melchior, M., Bolze, C., Fombonne, E., Surkan, P. J., Pryor, L., and Jauffret-Roustide, M. (2017). Early cannabis initiation and educational attainment: is the association causal? Data from the French TEMPO study. *International Journal of Epidemiology*, 46(5):1641–1650.

Moffitt, R. (1993). Identification and estimation of dynamic models with a time series of repeated cross-sections. *Journal of Econometrics*, 59(1-2):99–123.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 49(6):1417–1426.

Nijman, T. and Verbeek, M. (1990). Estimation of time-dependent parameters in linear models using cross-sections, panel, or both. *Journal of Econometrics*, 46:333–346.

Onrust, S. A., Otten, R., Lammers, J., and Smit, F. (2016). School-based programmes to reduce and prevent substance use in different age groups: What works for whom? Systematic review and meta-regression analysis. *Clinical Psychology Review*, 44:45–59.

Orosová, O., Gajdošová, B., Bačíková-šléšková, M., Benka, J., and Bavoľár, J. (2020). Alcohol consumption among slovak schoolchildren: Evaluation of the effectiveness of the unplugged programme. *Adiktologie*, 20(3-4):89–96.

Pan, W. (2001). Akaike's information criterion in generalized estimating equations. *Biometrics*, 57(March):120–125.

Pearson, M. R., Hustad, J. T., Neighbors, C., Conner, B. T., and Bravo, A. J. (2018). Personality, marijuana norms, and marijuana outcomes among college students. *Addictive Behaviors*, 76:291–297.

Perry, C. L., Komro, K. A., Veblen-Mortenson, S., Bosma, L. M., Farbakhsh, K., Munson, K. A., Stigler, M. H., and Lytle, L. A. (2003). A randomized controlled trial of the middle and junior high school D.A.R.E. and D.A.R.E. Plus programs. *Archives of Pediatrics and Adolescent Medicine*, 157(2):178–184.

Poudel, A. and Gautam, S. (2017). Age of onset of substance use and psychosocial problems among individuals with substance use disorders. *BMC Psychiatry*, 17(1):1–7.

Register, C. A., Williams, D. R., and Grimes, P. W. (2001). Adolescent drug use and educational attainment. *Education Economics*, 9(1):1–18.

Roebuck, M. C., French, M. T., and Dennis, M. L. (2004). Adolescent marijuana use and school attendance. *Economics of Education Review*, 23(2):133–141.

Rosenbaum, D. P. (2007). Just Say No To D.a.R.E. *Criminology & Public Policy*, 6(4):815–824.

Schmidt, L., Weisner, C., and Wiley, J. (1998). Substance abuse and the course of welfare dependency. *American Journal of Public Health*, 88:1616–1622.

Sen, B., Agarwal, R., and Hofler, R. (2002). Teenage indulgence in cigarettes, alcohol and marijuana: Evidence of a ”gateway” effect. (March).

Shackleton, N., Jamal, F., Viner, R. M., Dickson, K., Patton, G., and Bonell, C. (2016). School-based interventions going beyond health education to promote adolescent health: Systematic review of reviews. *Journal of Adolescent Health*, 58(4):382–396.

Soole, D. W., Mazerolle, L., and Rombouts, S. (2008). School-based drug prevention programs: A review of what works. *Australian and New Zealand Journal of Criminology*, 41(2):259–286.

Steinberg, L. (2010). A dual systems model of adolescent risk-taking. *Developmental Psychobiology*, 52(3):216–224.

Stuyt, E. (2018). The problem with the current high potency THC marijuana from the perspective of an addiction psychiatrist. *Missouri Medicine*, 115(6):482–487.

Sweet, S. A. and Grace-Martin, K. (2008). *Data analysis with SPSS: A first course in applied statistics*. Boston, Mass: Allyn and Bacon.

Tanner-Smith, E. E., Durlak, J. A., and Marx, R. A. (2018). Empirically based mean effect size distributions for universal prevention programs targeting school-aged youth: A review of meta-analyses. *Prevention Science*, 19:1091–1101.

Tobler, N. S., Lessard, T., Marshall, D., Ochshorn, P., and Roona, M. (1999). Effectiveness of school-based drug prevention programs for marijuana use. *School Psychology International*, 20(1):105–137.

Tolan, P., Szapocznik, J., and Sambrano, S. (2007). *Preventing youth substance abuse*. American Psychological Association, Washington, DC.

Torres-Reyna, O. (2009). Getting started in factor analysis (using Stata 10).

Van Den Berg, G. J., Van Der Klaauw, B., and Van Ours, J. C. (2004). Punitive sanctions and the transition rate from welfare to work. *Journal of Labor Economics*, 22(1):211–241.

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

Verbeek, M. (1995). Pseudo panel data. In Màtyàs, L. and Sevestre, P., editors, *The Econometrics of Panel Data: hand book of the theory with applications*, number Second Revised Edition, chapter 11, pages 283–292. KLUWER ACADEMIC PUBLISHERS.

Verbeek, M. (2008). Pseudo-panels and repeated cross-sections. In Màtyàs, L. and Sevestre, P., editors, *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*, chapter 11, pages 369–384. Springer.

Verbeek, M. and Nijman, T. (1992). Can cohort data be treated as genuine panel data? *Empirical Economics*, 17:9–23.

Verweij, K. J., Huizink, A. C., Agrawal, A., Martin, N. G., and Lynskey, M. T. (2013). Is the relationship between early-onset cannabis use and educational attainment causal or due to common liability? *Drug and Alcohol Dependence*, 133:580–586.

Vigna-Taglianti, F., Vadrucci, S., Faggiano, F., Burkhart, G., Siliquini, R., and Galanti, M. R. (2009). Is universal prevention against youths' substance misuse really universal? Gender-specific effects in the EU-Dap school-based prevention trial. *Journal of Epidemiology and Community Health*, 63(9):722–728.

White, D. and Pitts, M. (1998). Educating young people about drugs: A systematic review. *Addiction*, 93(October 1997):1475–1487.

Williams, J. and Van Ours, J. C. (2020). Hazardous or not? Cannabis use and early labor market experiences of young men. *Health Economics*, 29:1148–1160.

Wilson, D. B., Gottfredson, D. C., and Najaka, S. S. (2001). School-based prevention of problem behaviors: a meta analysis. *J Quantiative Criminology*, 17(3):245–272.

Yamada, T., Kendix, M., and Yamada, T. (1996). The impact of alcohol consumption and marijuana use on high school graduation. *Health Economics*, 5:77–92.

Appendix A

General Information

A.1 Logit vs. Logistic Regression

In the generalized linear model, $E(y)$ is related to x and β by a link function g such as : $g(E(y)) = \beta^T x$, and logit and logistic regression are the same but describe the GLM in two ways.

1) Logit function is $logit(x) = log\frac{x}{1-x}$, it describes the GLM in term of link function and interprets the results in terms of predicted probabilities.

2) Logistic function is $f(x) = \frac{1}{1-\exp^{(-x)}}$, it describes the GLM in term of activation function and interpret coefficients in terms of odds ratios.

A.2 Hausman Test

In 1978 Hausman introduced Hausman test which compares two estimations that based on H0, one that is asymptotically efficient estimator, has zero asymptotic covariance from the consistent but inefficient estimator. However the H1 is alternative hypothesis that believe that only consistent estimator exist. The most well-know use of Hausman test is about comparing the random and fixed effect estimator and choosing the most consistent and efficient one. However another important use of this test is recognizing existence of sample selection bias with comparing unbalance and balance panel which is emphasised by Verbeek (2008). In 2019 Aït-Sahalia and Xiu summarized the general rule of Hausman test as below:

Table A.1: The general rule of Hausman test

	Estimator A	Estimator B
$H_0 : Null$	Consistent and efficient	Consistent
$H_0 : Alternative$	Inconsistent	Consistent

$$\xi = (\hat{\beta}_B - \hat{\beta}_A)\acute{}[\hat{V}\{\hat{\beta}_B\} - \hat{V}\{\hat{\beta}_A\}]^{-1}(\hat{\beta}_B - \hat{\beta}_A)$$
$$\begin{cases} \xi \to \chi^2 & Under H_0 \\ \xi \to \infty & Under H_1 \end{cases} \quad \text{(A.1)}$$

$\hat{V}$= Estimated covariance matrix

A = If H_0 cannot be rejected this estimation is consistent and efficient

B = If H_0 is rejected, this estimation is consistent

$H_0 = plim(\hat{\beta}_B - \hat{\beta}_A) = 0$

If $\xi < 0$ model fitted on the data fails to meet the asymptotic assumptions of the Hausman test. In this regard, changing the place of A and B solves the problem based on general form of Hausman test.[1]

[1]All Hausman test analysis for this study are available based on request.

A.3 Categorized Variables in Pseudo-Panel

Table A.2: Categorical variables in pseudo-panel
and their counterparts in pooled NSDUH

Variables	Pseudo-panel	Pooled NSDUH
	percentage	percentage
Peers use marijuana		
None of them	37	32
A few of them	42	46
Most of them	20	21
All of them	1	1
	Value	Value
Mean	0.9	0.9
Std. dev.	0.8	0.8
	percentage	percentage
The reported risk perception		
No risk	11	10
Slight risk	28	25
Moderate risk	32	33
Great risk	29	32
	Value	Value
Mean	2.9	2.9
Std. dev.	1.0	1.0

Appendix B

Information Regarding Chapter 1

In NSDUH (2002-2014) cross-sections database, income is coded in 7 groups (B.1 and I used consumer price index for all urban consumers published in the page 78 of CPI Detailed Report of Bureau of labor statistics 2016 to change the total family income to the total real family income.[1]

Table B.1: Codes and assumption for total family income in NSDUH (2002-2014)

Code	Range of income (Dollar)	Average of income (Dollar)
1	Less than 10,000 (Including Loss)	5000
2	10,000 - 19,999	14,999.5
3	20,000 - 29,999	24,999.5
4	30,000 - 39,999	34,999.5
5	40,000 - 49,999	44,999.5
6	50,000 - 74,999	62,499.5
7	75,000 or more	100,000

[1] More information regarding calculation is available upon request

Table B.2: Variables and their definitions

Explanation	Categories Categories	original codes	changed codes
Total real family income (Income)			
(Year 2000 is considered as the base year for all papers)			
	N/A		
At which age did you start using marijuana? (onset_M)			
	N/A		
The highest degree received (Edu)			
*The highest degree at age of 25 (NSLY97)			
*The highest degree at ages of 24-25 (NSDUH)			
	Less than High school diploma	0	0
	High school diploma and Higher	1	1
Do you like to test yourself by doing risky things? (Risk)			
	Never	1	1
	Seldom	2	2
	Sometimes/always	3	3
Race			
	Not Black-Not Hispanic	1	1
	Black	2	2
	Hispanic	3	3
Do you think using marijuana once a month is risky? (Risky_marijuana)			
	No risk	1	0
	Slight risk	2	0.33
	Moderate risk	3	0.67
	Great risk	4	1
What do your parents feel about you using marijuana monthly? (Parents_feel_use)			
	Neither approve or disapproved	1	0
	Somewhat Disapprove	2	0.5
	Strongly Disapprove	3	1
Did you have any drug education in school during This year? (infoclass)			
Ever used marijuana? (MJEVER)			
Have you used marijuana during this year? (MRJYR)			
Have you drank Alcohol during this year? (ALCYR)			
Have you smoked cigarette during this year? (CIGYR)			
Have you ever arrested or booked for breaking the law? (breakingLaw)			
Are you enrolled at any school right now? (Enrolled)	No	0	0
	Yes	1	1
HOW often have your parents said you did a good job during this year? (Parents_goodjob)			
HOW often have your parents said they are proud of you for what you had done? (Parents_proud)			
HOW often have your parents checked your homework during this year? (Parents_check)			
HOW often have your parents helped you doing the homework during this year? (Parents_help)			
	Never	1	0
	Seldom	2	0.33
	Sometimes	3	0.67
	always	4	1
Have you ever get into a serious fight at school or work? (Fight)			
Have you ever taken part in a group fight? (Group_fight)			
Have you ever argued with at least one of your parents? (Parents_fight)			
	0 time	1	0
	1 to 2 times	2	0.25
	3 to 5 times	3	0.5
	6 to 9 times	4	0.75
	10 times and more	5	1
Describe your overall health. (HEALTH)			
	Poor	1	0
	Fair	2	0.25
	Good	3	0.5
	Very good	4	0.75
	Excellent	5	1
Do you live in metropolitan city? (COUTYP2)			
	No metropolitan	0	0
	large/small metropolitan	1	1
How many students in your grade do you know who use marijuana? (peers_marijuana)			
	None of them	0	0
	A few of them	1	0.33
	Most of them	2	0.67
	All of them	3	1

I convert the original codes to new codes between 0 and 1 by $\frac{X-Min}{Max-Min}$ in cross-sectionals NSDUH (2002-2014) and after that I create NSDUH pseudo-panel.

Table B.3: Details on sample selection of NSDUH cross-sectionals (2002-2014)

Condition	
NSDUH 2002-2014 (Age 12 and older)	722,653
Individuals at ages 12 to 17 who answered the question regarding risk attitude (one of the factors for create Pseudo-panel)	217,814
All cohorts have to come to pseudo- panel when they are 12 years old and stay at least for 2 periods) (Table 1.1)	173,575
Individuals who registered to any schools (only 0.9% in total did not registered in any school)	171,989
Individuals have not started using marijuana before 13 year old (I checked age of start using marijuana and used marijuana$_{(t)}$ for 12 years old individuals	168,209

Table B.4: Details on sample selection of NSDUH pseudo-panel

Condition	
Cohort between ages 13 to 17	216
Remaining Cohorts$\times$ number of Years of staying in the pseudo-panel	900

Table B.5: Details on sample selection of NLSY97 panel data

Condition	
Number of Participants (Age 12 to 16) in the first wave (1997)	8,984
Cross-sectional samples in the first wave (1997)	6,748
Number of Participants (age 12) in the first wave	1,333
12 years participants who had not used marijuana at 1997	968
Remaining Participants$\times$ number of Yeas of staying in the NLSY97 panel consistent with pseudo-panel (Table 1.7)	4,789

Table B.6: Fixed effect and random effect linear models weighted pseudo-panel

	FE	RE	FE	RE	FE	RE	FE	RE
The reported risk perception								
No risk	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*
Slight risk	−0.196**	−0.300***	−0.180*	−0.199**	−0.194**	−0.218***	−0.194**	−0.216***
	[0.075]	[0.069]	[0.074]	[0.066]	[0.071]	[0.061]	[0.071]	[0.061]
Moderate risk	−0.198*	−0.322***	−0.181*	−0.212**	−0.209**	−0.228***	−0.209**	−0.199**
	[0.082]	[0.070]	[0.081]	[0.066]	[0.080]	[0.062]	[0.080]	[0.062]
Great risk	−0.223**	−0.307***	−0.198*	−0.178**	−0.255**	−0.251***	−0.255**	−0.223***
	[0.083]	[0.071]	[0.080]	[0.067]	[0.081]	[0.062]	[0.081]	[0.062]
Drug info$_{(t-1)}$=1	0.014	0.023	0.011	0.017	0.006	0.014	0.006	0.016
	[0.016]	[0.015]	[0.016]	[0.014]	[0.017]	[0.013]	[0.017]	[0.013]
Age								
13	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*
14	0.021	−0.011	0.018	−0.004	0	−0.031	0	−0.034
	[0.036]	[0.033]	[0.034]	[0.030]	[0.030]	[0.028]	[0.030]	[0.028]
15	0.033	−0.002	0.025	−0.008	0.001	−0.024	0.001	−0.018
	[0.021]	[0.021]	[0.018]	[0.019]	[0.018]	[0.018]	[0.018]	[0.018]
16	0.102***	0.049*	0.085***	0.016	0.055**	0	0.055**	0.019
	[0.020]	[0.020]	[0.017]	[0.019]	[0.017]	[0.017]	[0.017]	[0.018]
17	0.144***	0.077***	0.118***	0.014	0.083***	0	0.083***	0.027
	[0.017]	[0.019]	[0.016]	[0.018]	[0.015]	[0.017]	[0.015]	[0.017]
male		0.009		0.014*		0.019**		0.01
		[0.008]		[0.007]		[0.006]		[0.007]
Race								
Not-Black, not-Hispanic		*Reference*		*Reference*		*Reference*		*Reference*
Black		−0.064*		−0.035		−0.038		−0.028
		[0.026]		[0.024]		[0.022]		[0.022]
Hispanic		−0.094***		−0.056**		−0.070***		−0.057**
		[0.023]		[0.021]		[0.020]		[0.020]
Caring family$_{(t-1)}$	−0.111***	−0.187***	−0.086**	−0.103***	−0.071*	−0.061**	−0.071*	−0.046*
	[0.030]	[0.024]	[0.030]	[0.023]	[0.030]	[0.022]	[0.030]	[0.022]
Peers use marijuana								
Non of them	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*
A few of them	−0.046	−0.052	−0.044	−0.047	−0.036	−0.045	−0.036	−0.041
	[0.026]	[0.031]	[0.023]	[0.029]	[0.019]	[0.027]	[0.019]	[0.027]
Most of them	−0.011	−0.007	−0.01	−0.013	−0.006	−0.015	−0.006	−0.014
	[0.023]	[0.030]	[0.021]	[0.028]	[0.016]	[0.026]	[0.016]	[0.026]
All of them	0.062*	0.075*	0.056	0.037	0.055*	0.027	0.055*	0.025
	[0.032]	[0.035]	[0.030]	[0.033]	[0.028]	[0.031]	[0.028]	[0.030]
Real family income	−0.009	−0.021	−0.007	−0.011	−0.014	−0.027**	−0.014	−0.023*
	[0.014]	[0.011]	[0.013]	[0.011]	[0.013]	[0.010]	[0.013]	[0.010]
Live in metropolitan city=1 (>0.75)	0.011	0.012	0.011	0.01	0.015	0.016**	0.015	0.016**
	[0.007]	[0.006]	[0.007]	[0.006]	[0.008]	[0.005]	[0.008]	[0.005]
The reported risk perception#Drug info$_{(t-1)}$=1								
No risk#Drug info$_{(t-1)}$=1	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*
Slight risk#Drug info$_{(t-1)}$=1	0.001	−0.012	0.003	−0.013	0.007	−0.011	0.007	−0.014
	[0.018]	[0.016]	[0.019]	[0.015]	[0.019]	[0.014]	[0.019]	[0.014]
Moderate risk#Drug info$_{(t-1)}$=1	−0.005	−0.025	−0.001	−0.02	0.004	−0.017	0.004	−0.02
	[0.018]	[0.018]	[0.018]	[0.016]	[0.019]	[0.015]	[0.019]	[0.015]
Great risk#Drug info$_{(t-1)}$=1	−0.007	−0.019	−0.003	−0.017	−0.002	−0.013	−0.002	−0.016
	[0.021]	[0.020]	[0.021]	[0.018]	[0.021]	[0.017]	[0.021]	[0.017]

Continued on next page

Table B.7: Fixed effect and random effect linear models weighted pseudo-panel (continued from previous page)

Continued from previous page	FE	RE	FE	RE	FE	RE	FE	RE
The reported risk perception#Age								
No risk#13	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*
Slight risk#14	−0.003	0.039	−0.003	0.029	0.01	0.046	0.01	0.06
	[0.040]	[0.036]	[0.039]	[0.033]	[0.034]	[0.031]	[0.034]	[0.030]
Slight risk#15	0.053	0.076**	0.052*	0.063**	0.063**	0.063**	0.063**	0.079***
	[0.028]	[0.025]	[0.026]	[0.023]	[0.024]	[0.022]	[0.024]	[0.022]
Slight risk#16	0.021	0.041	0.018	0.029	0.026	0.026	0.026	0.041
	[0.028]	[0.025]	[0.026]	[0.023]	[0.024]	[0.021]	[0.024]	[0.021]
Slight risk#17	0.014	0.036	0.015	0.037	0.022	0.029	0.022	0.042*
	[0.024]	[0.024]	[0.022]	[0.022]	[0.020]	[0.020]	[0.020]	[0.020]
Moderate risk#14	0.033	0.036	0.029	0.017	0.04	0.04	0.04	0.056
	[0.037]	[0.034]	[0.035]	[0.031]	[0.031]	[0.029]	[0.031]	[0.029]
Moderate risk#15	0.076**	0.059**	0.070**	0.035	0.079***	0.045*	0.079***	0.061**
	[0.024]	[0.023]	[0.021]	[0.021]	[0.020]	[0.019]	[0.020]	[0.019]
Moderate risk#16	0.044*	0.027	0.040*	0.013	0.043*	0.011	0.043*	0.024
	[0.022]	[0.022]	[0.020]	[0.020]	[0.018]	[0.019]	[0.018]	[0.019]
Moderate risk#17	0.016	0.004	0.016	0.007	0.021	0.004	0.021	0.011
	[0.019]	[0.021]	[0.016]	[0.020]	[0.015]	[0.019]	[0.015]	[0.018]
Great risk#14	0.024	0.033	0.022	0.014	0.031	0.041	0.031	0.052
	[0.035]	[0.033]	[0.034]	[0.030]	[0.030]	[0.028]	[0.030]	[0.028]
Great risk#15	0.048**	0.033	0.044**	0.013	0.054***	0.029	0.054***	0.037*
	[0.018]	[0.020]	[0.016]	[0.019]	[0.015]	[0.017]	[0.015]	[0.017]
Great risk#16	0.033*	0.023	0.032*	0.017	0.037**	0.025	0.037**	0.027
	[0.015]	[0.019]	[0.014]	[0.018]	[0.013]	[0.017]	[0.013]	[0.017]
Great risk#17	0	0	0	0	0	0	0	0
	[.]	[.]	[.]	[.]	[.]	[.]	[.]	[.]
The reported risk perception#Gender								
No risk#Male	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*
Slight risk#Male	−0.003	−0.005	−0.005	−0.013	−0.007	−0.009	−0.007	−0.005
	[0.015]	[0.009]	[0.015]	[0.008]	[0.015]	[0.007]	[0.015]	[0.007]
Moderate risk#Male	−0.011	−0.012	−0.012	−0.014	−0.017	−0.012	−0.017	−0.006
	[0.015]	[0.009]	[0.014]	[0.008]	[0.014]	[0.007]	[0.014]	[0.007]
Great risk#Male	−0.013	−0.011	−0.015	−0.017*	−0.025	−0.019*	−0.025	−0.01
	[0.015]	[0.009]	[0.014]	[0.008]	[0.014]	[0.008]	[0.014]	[0.008]
The reported risk perception#Race								
No risk#not Black-not Hispanic	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*
Slight risk#Black	0.039	0.088**	0.033	0.05	0.021	0.028	0.021	0.026
	[0.034]	[0.029]	[0.033]	[0.027]	[0.032]	[0.025]	[0.032]	[0.025]
Slight risk#Hispanic	0.051	0.092***	0.042	0.055*	0.041	0.050*	0.041	0.048*
	[0.031]	[0.026]	[0.029]	[0.024]	[0.028]	[0.023]	[0.028]	[0.022]
Moderate risk#Black	0.043	0.096**	0.036	0.057*	0.032	0.034	0.032	0.022
	[0.035]	[0.029]	[0.034]	[0.027]	[0.033]	[0.025]	[0.033]	[0.025]
Moderate risk#Hispanic	0.06	0.106***	0.048	0.064**	0.054	0.055*	0.054	0.045*
	[0.032]	[0.026]	[0.030]	[0.024]	[0.030]	[0.023]	[0.030]	[0.023]
Great risk#Black	0.069	0.088**	0.058	0.032	0.064	0.035	0.064	0.026
	[0.037]	[0.032]	[0.035]	[0.030]	[0.034]	[0.028]	[0.034]	[0.027]
Great risk#Hispanic	0.086*	0.104***	0.069*	0.046	0.087**	0.060*	0.087**	0.051*
	[0.034]	[0.028]	[0.030]	[0.026]	[0.031]	[0.024]	[0.031]	[0.024]
The reported risk perception#caring family								
No risk#caring family	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*
Slight risk#caring family	−0.008	−0.016	−0.017	−0.031	−0.005	−0.021	−0.005	−0.018
	[0.035]	[0.027]	[0.036]	[0.025]	[0.035]	[0.023]	[0.035]	[0.023]
Moderate risk#caring family	0.103**	0.088**	0.079*	0.024	0.077*	0.023	0.077*	0.027
	[0.035]	[0.028]	[0.035]	[0.026]	[0.035]	[0.024]	[0.035]	[0.024]
Great risk#caring family	0.234***	0.172***	0.188***	0.054	0.166***	0.047	0.166***	0.066*
	[0.034]	[0.031]	[0.035]	[0.029]	[0.035]	[0.027]	[0.035]	[0.028]

Continued on next page

Table B.8: Fixed effect and random effect linear models weighted pseudo-panel (continued from previous page)

	FE	RE	FE	RE	FE	RE	FE	RE
The reported risk perception#Peers use marijuana								
No risk#Non of them	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*
Slight risk#A few of them	0.056	0.033	0.055	0.034	0.051*	0.044	0.051*	0.038
	[0.030]	[0.034]	[0.028]	[0.032]	[0.024]	[0.030]	[0.024]	[0.030]
Slight risk#Most of them	0.046	0.021	0.047	0.026	0.050*	0.044	0.050*	0.037
	[0.029]	[0.034]	[0.028]	[0.032]	[0.022]	[0.030]	[0.022]	[0.029]
Slight risk#All of them	0.022	−0.02	0.026	0.002	0.041	0.051	0.041	0.044
	[0.054]	[0.046]	[0.055]	[0.044]	[0.054]	[0.041]	[0.054]	[0.040]
Moderate risk#A few of them	0.045	0.063*	0.047	0.065*	0.041	0.061*	0.041	0.05
	[0.027]	[0.032]	[0.025]	[0.030]	[0.021]	[0.028]	[0.021]	[0.027]
Moderate risk#Most of them	0.025	0.037	0.026	0.036	0.024	0.043	0.024	0.032
	[0.026]	[0.032]	[0.024]	[0.031]	[0.021]	[0.028]	[0.021]	[0.028]
Moderate risk#All of them	0	0	0	0	0	0	0	0
	[.]	[.]	[.]	[.]	[.]	[.]	[.]	[.]
Great risk#A few of them	0.041	0.057	0.04	0.056*	0.033	0.048	0.033	0.043
	[0.025]	[0.029]	[0.023]	[0.028]	[0.018]	[0.026]	[0.018]	[0.025]
Great risk#Most of them	0	0	0	0	0	0	0	0
	[.]	[.]	[.]	[.]	[.]	[.]	[.]	[.]
Great risk#All of them	0	0	0	0	0	0	0	0
	[.]	[.]	[.]	[.]	[.]	[.]	[.]	[.]
The reported risk perception#Real income								
No risk#Real income	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*
Slight risk#Real income	0.018	0.039**	0.015	0.024*	0.017	0.024*	0.017	0.023*
	[0.015]	[0.013]	[0.014]	[0.012]	[0.014]	[0.011]	[0.014]	[0.011]
Moderate risk#Real income	0.012	0.035**	0.01	0.021	0.016	0.023*	0.016	0.017
	[0.016]	[0.013]	[0.015]	[0.012]	[0.015]	[0.011]	[0.015]	[0.011]
Great risk#Real income	0.012	0.026*	0.009	0.011	0.024	0.025*	0.024	0.02
	[0.016]	[0.013]	[0.015]	[0.012]	[0.015]	[0.011]	[0.015]	[0.011]
The reported risk perception#Live in metropolitan city								
No risk#Live in metropolitan city=1 (>0.75)	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*
Slight risk#Live in metropolitan city=1 (>0.75)	−0.011	−0.015	−0.012	−0.012	−0.013	−0.013	−0.013	−0.012
	[0.008]	[0.008]	[0.008]	[0.007]	[0.009]	[0.007]	[0.009]	[0.007]
Moderate risk#Live in metropolitan city=1 (>0.75)	−0.01	−0.011	−0.011	−0.014	−0.014	−0.017*	−0.014	−0.016*
	[0.008]	[0.008]	[0.009]	[0.007]	[0.009]	[0.007]	[0.009]	[0.007]
Great risk#Live in metropolitan city=1 (>0.75)	−0.018*	−0.015	−0.018*	−0.011	−0.021*	−0.017*	−0.021*	−0.018*
	[0.009]	[0.008]	[0.009]	[0.008]	[0.009]	[0.007]	[0.009]	[0.007]
Marijuana use$_{(t-1)}$			0.146**	0.513***	0.011	0.242***	0.011	0.239***
			[0.047]	[0.031]	[0.051]	[0.038]	[0.051]	[0.038]
Drink alcohol$_{(t-1)}$					0.170***	0.194***	0.170***	0.148***
					[0.028]	[0.022]	[0.028]	[0.024]
Smoke cigarette$_{(t-1)}$					0.034	0.080*	0.034	0.075*
					[0.052]	[0.036]	[0.052]	[0.036]
Like to do risky things								
Never								*Reference*
Seldom								0.010*
								[0.004]
Sometimes-Always								0.028***
								[0.006]
Constant	0.171*	0.295***	0.163*	0.205***	0.180**	0.264***	0.180**	0.229***
	[0.067]	[0.063]	[0.066]	[0.060]	[0.064]	[0.056]	[0.064]	[0.055]
N	900	900	900	900	900	900	900	900
AIC	−4200	−3600	−4200	−3800	−4300	−3900	−4300	−3900
BIC	−3900	−3300	−3900	−3500	−4000	−3600	−4000	−3600
sigma_u	0.043	0.016	0.037	0	0.036	0	0.036	0
sigma_e	0.027	0.027	0.026	0.028	0.026	0.026	0.026	0.025
rho	0.718	0.263	0.666	0	0.664	0	0.664	0
R-squared Overall	0.775		0.812		0.826		0.826	
log likelihood		1869		1955		2023		2037

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

The standard deviation is inside the bracket [...]

FE linear weighted regression: xtreg [pweight=we], fe vce(robust)

RE linear weighted regression:xtreg [iweight=we], re mle

Table B.9: Marijuana use and school-based drug education
pooled weighted OLS pseudo-panel and weighted fixed effect linear pseudo-panel (2002-2014)

Marijuana use(t)	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Pooled	*FE*	*Pooled*	*FE*	*Pooled*	*FE*	*Pooled*	*FE*	*Pooled*	*FE*	*Pooled*	*FE*
Drug info=1	0.213**	0.097	0.173**	0.055	−0.007	0.051	−0.047	−0.006	0.013	0.113*	0.03	0.110*
	[0.078]	[0.064]	[0.059]	[0.056]	[0.049]	[0.056]	[0.047]	[0.051]	[0.035]	[0.044]	[0.037]	[0.044]
Male	0.043*		0.008		0.017*		0.026***		0.018**		0.019***	
	[0.018]		[0.011]		[0.008]		[0.007]		[0.005]		[0.005]	
Race												
Non-Black, Non-Hispanic	*Reference*		*Reference*		*Reference*		*Reference*		*Reference*		*Reference*	
Black	0.035*		0.02		0.006		−0.019*		0.035***		0.022*	
	[0.016]		[0.011]		[0.008]		[0.008]		[0.007]		[0.010]	
Hispanic	0.023		0.017		−0.021*		−0.031***		0.005		−0.008	
	[0.017]		[0.011]		[0.009]		[0.009]		[0.006]		[0.009]	
Age	0.073***	0.066***	0.045***	0.048***	0.024***	0.042***	0.021***	0.037***	0.012***	0.016***	0.011***	0.017***
	[0.004]	[0.003]	[0.003]	[0.003]	[0.003]	[0.004]	[0.003]	[0.004]	[0.002]	[0.003]	[0.003]	[0.003]
Male# Drug info=1	−0.044*	−0.042***	−0.031*	−0.025**	−0.027**	−0.023**	−0.030***	−0.029***	−0.014*	−0.016**	−0.014*	−0.015*
	[0.018]	[0.011]	[0.012]	[0.009]	[0.008]	[0.008]	[0.008]	[0.007]	[0.006]	[0.006]	[0.006]	[0.006]
Race#drug info interaction												
Non-Black, non-Hispanic	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*
Black# Drug info=1	−0.043*	−0.013	−0.033**	−0.013	−0.012	−0.011	0.002	0.002	−0.024***	−0.034***	−0.026***	−0.034***
	[0.017]	[0.012]	[0.012]	[0.009]	[0.009]	[0.009]	[0.009]	[0.008]	[0.007]	[0.007]	[0.007]	[0.007]
Hispanic# Drug info=1	−0.023	−0.008	−0.018	−0.008	−0.014	−0.007	−0.005	0.001	−0.006	−0.005	−0.007	−0.004
	[0.019]	[0.013]	[0.012]	[0.011]	[0.009]	[0.011]	[0.009]	[0.010]	[0.007]	[0.008]	[0.007]	[0.008]
Age# Drug info=1	−0.009	−0.004	−0.008*	−0.002	0.002	−0.002	0.004	0.002	0	−0.006*	−0.001	−0.005*
	[0.005]	[0.004]	[0.004]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.002]	[0.003]	[0.002]	[0.003]
The reported risk perception												
No risk			*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*
Slight risk			−0.092***	−0.069***	−0.067***	−0.068***	−0.053***	−0.058***	−0.052***	−0.042***	−0.052***	−0.042***
			[0.009]	[0.007]	[0.007]	[0.007]	[0.006]	[0.005]	[0.004]	[0.004]	[0.004]	[0.004]
Moderate risk			−0.149***	−0.103***	−0.099***	−0.102***	−0.084***	−0.091***	−0.080***	−0.061***	−0.079***	−0.060***
			[0.009]	[0.009]	[0.007]	[0.009]	[0.006]	[0.007]	[0.005]	[0.006]	[0.005]	[0.006]
Great risk			−0.175***	−0.102***	−0.097***	−0.098***	−0.088***	−0.099***	−0.089***	−0.059***	−0.087***	−0.060***
			[0.009]	[0.012]	[0.008]	[0.012]	[0.007]	[0.009]	[0.005]	[0.007]	[0.006]	[0.007]
Caring family					−0.190***	−0.060***	−0.168***	−0.056***	−0.020*	0.001	−0.023*	0
					[0.009]	[0.017]	[0.009]	[0.016]	[0.009]	[0.013]	[0.010]	[0.013]
Peers use marijuana												
None of them							*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*
A few of them							−0.010*	−0.017***	−0.011***	−0.008**	−0.010***	−0.009**
							[0.004]	[0.004]	[0.003]	[0.003]	[0.003]	[0.003]
Most of them							0.037***	0.023**	0.027***	0.027***	0.028***	0.027***
							[0.008]	[0.008]	[0.006]	[0.006]	[0.006]	[0.006]
All of them							0.102***	0.085***	0.097***	0.082***	0.097***	0.083***
							[0.023]	[0.023]	[0.021]	[0.023]	[0.021]	[0.023]
Drink alcohol									0.150***	0.225***	0.168***	0.225***
									[0.019]	[0.024]	[0.021]	[0.024]
Smoke cigarette									0.330***	0.379***	0.330***	0.379***
									[0.029]	[0.034]	[0.029]	[0.033]
Real family income											−0.007*	0.006
											[0.004]	[0.005]
Live in metropolitan city=1											−0.001	−0.002
											[0.003]	[0.002]
Like to do risky things												
Never											*Reference*	
Seldom											0.00	
											[0.004]	
Sometimes / always											−0.003	
											[0.006]	
Constant	−1.034***	−0.863***	−0.472***	−0.524***	−0.176***	−0.433***	−0.145**	−0.352***	−0.103**	−0.208***	−0.054	−0.236***
	[0.071]	[0.055]	[0.057]	[0.058]	[0.049]	[0.062]	[0.047]	[0.057]	[0.036]	[0.050]	[0.044]	[0.054]
N	900	900	900	900	900	900	900	900	900	900	900	900
AIC	−2300	−3400	−2800	−3700	−3300	−3700	−3400	−3900	−4000	−4300	−4000	−4300
BIC	−2300	−3400	−2800	−3700	−3200	−3700	−3300	−3800	−3900	−4200	−3900	−4200
R-squared	0.629	0.835	0.794	0.883	0.875	0.885	0.893	0.91	0.94	0.94	0.94	0.94

$^{*}p < 0.05, ^{**}p < 0.01, ^{***}p < 0.001$
The standard deviation is inside the bracket [...]
This table shows simple comparison between pooled and FE models, but model 6 (FE) is not the final model, because the time lags are not considered here.

Table B.10: Marijuana use and school-based drug education (2002-2014) pooled weighted OLS pseudo-panel and pooled OLS NSDUH

Marijuana use$_{(t)}$	Pooled OLS NSDUH	Pooled weighted OLS Pseudo-panel
Drug info=1	0.049^{*}	0.03
	[0.020]	[0.037]
Male	0.005	0.019^{***}
	[0.003]	[0.005]
Race		
Non-Black, Non-Hispanic	*Reference*	*Reference*
Black	0.019^{***}	0.022^{*}
	[0.005]	[0.010]
Hispanic	0.005	-0.008
	[0.004]	[0.009]
Age	0.012^{***}	0.011^{***}
	[0.001]	[0.003]
Male # Drug info=1	0.002	-0.014^{*}
	[0.004]	[0.006]
Race drug info interaction		
Non-Black, non-Hispanic	*Reference*	*Reference*
Black# Drug info=1	-0.013^{*}	-0.026^{***}
	[0.005]	[0.007]
Hispanic# Drug info=1	-0.009	-0.007
	[0.005]	[0.007]
Age# Drug info=1	-0.003^{**}	-0.001
	[0.001]	[0.002]
The reported risk perception		
No risk	*Reference*	*Reference*
Slight risk	-0.159^{***}	-0.052^{***}
	[0.003]	[0.004]
Moderate risk	-0.212^{***}	-0.079^{***}
	[0.003]	[0.005]
Great risk	-0.218^{***}	-0.087^{***}
	[0.003]	[0.006]
Caring family	-0.005^{***}	-0.023^{*}
	[0.001]	[0.010]
Peers use marijuana		
None of them	*Reference*	*Reference*
A few of them	-0.001	-0.010^{***}
	[0.002]	[0.003]
Most of them	0.104^{***}	0.028^{***}
	[0.003]	[0.006]
All of them	0.193^{***}	0.097^{***}
	[0.007]	[0.021]
Drink alcohol	0.153^{***}	0.168^{***}
	[0.002]	[0.021]
Smoke cigarette	0.296^{***}	0.330^{***}
	[0.003]	[0.029]
Real family income	0	-0.007^{*}
	[0.000]	[0.004]
Live in metropolitan city=1	0.016^{***}	-0.001
	[0.002]	[0.003]
Like to do risky things		
Never	*Reference*	*Reference*
Seldom	0.006^{**}	0
	[0.002]	[0.004]
Sometimes/always	0.013^{***}	-0.003
	[0.002]	[0.006]
Constant	-0.018	-0.054
	[0.018]	[0.044]
N	120000	900
AIC	22000	-4000
BIC	23000	-3900
R-squared	0.381	0.944

$^{*} p < 0.05, ^{**} p < 0.01, ^{***} p < 0.001$

The standard deviation is inside the bracket [...]

Table B.11: Fixed effect linear model, NSDUH weighted pseudo-panel investigating effects of the school-based drug education lags on marijuana use

	Final model one-year lag	two years lags	three years lags
Drug info$_{(t-1)}$=1	0.042	0.055	−0.041
	[0.059]	[0.102]	[0.178]
Age	0.013***	0.029***	0.047**
	[0.004]	[0.008]	[0.015]
Male # Drug info$_{(t-1)}$=1	−0.013	−0.01	−0.003
	[0.010]	[0.011]	[0.012]
Non-Black, non-Hispanic # Drug info$_{(t-1)}$=1	*Reference*	*Reference*	*Reference*
Black# Drug info$_{(t-1)}$=1	−0.016	−0.015	−0.021
	[0.013]	[0.013]	[0.013]
Hispanic# Drug info$_{(t-1)}$=1	−0.015	−0.012	−0.014
	[0.013]	[0.015]	[0.017]
Age# Drug info$_{(t-1)}$=1	−0.001	−0.002	0.004
	[0.003]	[0.006]	[0.011]
The reported risk perception			
No risk	*Reference*	*Reference*	*Reference*
Slight risk	−0.043***	−0.040***	−0.034***
	[0.005]	[0.005]	[0.007]
Moderate risk	−0.069***	−0.059***	−0.039***
	[0.006]	[0.007]	[0.010]
Great risk	−0.076***	−0.055***	−0.026
	[0.007]	[0.009]	[0.016]
Caring family$_{(t-1)}$	−0.022	−0.032*	−0.033
	[0.012]	[0.015]	[0.020]
Peers use marijuana			
None of them	*Reference*	*Reference*	*Reference*
A few of them	0.007*	−0.001	−0.042*
	[0.003]	[0.006]	[0.019]
Most of them	0.038***	0.028**	−0.021
	[0.007]	[0.008]	[0.021]
All of them	0.092***	0.089***	0.041
	[0.022]	[0.021]	[0.028]
Marijuana use$_{(t-1)}$	0.184***	0.063	−0.092
	[0.047]	[0.055]	[0.064]
Drink alcohol$_{(t-1)}$	0.212***	0.168***	0.124*
	[0.029]	[0.039]	[0.052]
Smoke cigarette$_{(t-1)}$	0.002	0.003	−0.064
	[0.055]	[0.061]	[0.076]
Real family income	0.002	0.003	0.001
	[0.005]	[0.007]	[0.008]
Live in metropolitan city=1	0	0.003	0.004
	[0.003]	[0.004]	[0.005]
Drug info$_{(t-2)}$=1		0.01	−0.062
		[0.117]	[0.222]
Male # Drug info$_{(t-2)}$=1		0.022	0.045
		[0.020]	[0.025]
Non-Black, non-Hispanic # Drug info$_{(t-2)}$=1		*Reference*	*Reference*
Black# Drug info$_{(t-2)}$=1		0.003	−0.016
		[0.021]	[0.027]
Hispanic# Drug info$_{(t-2)}$=1		−0.026	−0.024
		[0.022]	[0.027]
Age# Drug info$_{(t-2)}$=1		−0.001	0.002
		[0.007]	[0.013]
Drug info$_{(t-3)}$=1			0.06
			[0.190]
Male # Drug info$_{(t-3)}$=1			0.062**
			[0.023]
Non-Black, non-Hispanic # Drug info$_{(t-3)}$=1			*Reference*
Black# Drug info$_{(t-3)}$=1			−0.005
			[0.020]
Hispanic# Drug info$_{(t-3)}$=1			0
			[.]
Age# Drug info$_{(t-3)}$=1			−0.007
			[0.013]
Constant	−0.111	−0.341*	−0.519*
	[0.065]	[0.131]	[0.236]
N	900	684	486
AIC	−4100	−3100	−2200
BIC	−4000	−3000	−2100
sigma_u	0.023	0.036	0.076
sigma_e	0.028	0.03	0.031
rho	0.388	0.581	0.856
R-squared Overall	0.877	0.81	0.389

$^{*}p < 0.05, ^{**}p < 0.01, ^{***}p < 0.001$
The standard deviation is inside the bracket [...]

Table B.12: Fixed effect linear model, NSDUH weighted pseudo-panel investigating effects of the school-based drug education lags on risk of marijuana perception

	Final model one-year lag	two years lags	three years lags
Drug info$_{(t-1)}$=1	1.930**	2.841*	−0.865
	[0.687]	[1.195]	[2.401]
Age	−0.327***	−0.202*	−0.178
	[0.048]	[0.087]	[0.164]
Male # Drug info$_{(t-1)}$=1	−0.048	−0.155	−0.058
	[0.116]	[0.136]	[0.136]
Black# Drug info$_{(t-1)}$=1	−0.336*	−0.294	−0.335
	[0.164]	[0.185]	[0.191]
Hispanic# Drug info$_{(t-1)}$=1	−0.389*	−0.378	−0.402*
	[0.184]	[0.194]	[0.200]
Age# Drug info$_{(t-1)}$=1	−0.109**	−0.157*	0.071
	[0.041]	[0.071]	[0.142]
Caring family$_{(t-1)}$	0.549*	0.474*	0.619*
	[0.220]	[0.215]	[0.287]
Peers use marijuana			
Non of them	*reference*	*reference*	*reference*
A few of them	−0.116	−0.137	0.017
	[0.062]	[0.139]	[0.224]
Most of them	−0.071	−0.014	0.21
	[0.113]	[0.164]	[0.243]
All of them	−0.442*	−0.303	−0.194
	[0.185]	[0.230]	[0.311]
Real family income	−0.012	−0.061	0.038
	[0.081]	[0.093]	[0.132]
Live in metropolitan city=1	−0.028	−0.049	0
	[0.057]	[0.059]	[0.078]
Drug info$_{(t-2)}$=1		1.784	2.19
		[1.363]	[2.243]
Male # Drug info$_{(t-2)}$=1		−0.001	−0.097
		[0.193]	[0.259]
Black# Drug info$_{(t-2)}$=1		−0.385	−0.487
		[0.237]	[0.275]
Hispanic# Drug info$_{(t-2)}$=1		−0.225	−0.51
		[0.265]	[0.370]
Age# Drug info$_{(t-2)}$=1		−0.101	−0.113
		[0.081]	[0.135]
Drug info$_{(t-3)}$=1			1.775
			[2.538]
Male # Drug info$_{(t-3)}$=1			0.072
			[0.246]
Black# Drug info$_{(t-3)}$=1			0.113
			[0.245]
Hispanic# Drug info$_{(t-3)}$=1			0
			[.]
Age# Drug info$_{(t-3)}$=1			−0.132
			[0.163]
Constant	6.419***	4.498**	3.695
	[0.852]	[1.423]	[2.524]
N	900	684	486
AIC	940.311	567.943	280.2
BIC	997.94	644.919	363.925
sigma_u	0.685	0.739	0.839
sigma_e	0.466	0.432	0.405
rho	0.684	0.746	0.811
R-squared Overall	0.395	0.266	0.068

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

The standard deviation is inside the bracket [...]

Table B.13: Fixed effect linear model, NSDUH weighted pseudo-panel investigating interaction between age and peers marijuana use

	Final model	Model with age#peers marijuana use
Drug info$_{(t-1)}$=1	0.042	0.036
	[0.059]	[0.058]
Age	0.013***	0.016**
	[0.004]	[0.005]
Male # Drug info$_{(t-1)}$=1	−0.013	−0.014
	[0.010]	[0.010]
Black# Drug info$_{(t-1)}$=1	−0.016	−0.015
	[0.013]	[0.013]
Hispanic# Drug info$_{(t-1)}$=1	−0.015	−0.014
	[0.013]	[0.013]
Age# Drug info$_{(t-1)}$=1	−0.001	0
	[0.003]	[0.003]
The reported risk perception		
No risk	*Reference*	*Reference*
Slight risk	−0.043***	−0.042***
	[0.005]	[0.005]
Moderate risk	−0.069***	−0.069***
	[0.006]	[0.006]
Great risk	−0.076***	−0.075***
	[0.007]	[0.007]
Caring family$_{(t-1)}$	−0.022	−0.021
	[0.012]	[0.012]
Peers use marijuana		
Non of them	*Reference*	*Reference*
A few of them	0.007*	0.049
	[0.003]	[0.052]
Most of them	0.038***	0.122
	[0.007]	[0.080]
All of them	0.092***	−0.396
	[0.022]	[0.348]
Marijuana use$_{(t-1)}$	0.184***	0.190***
	[0.047]	[0.051]
Drink alcohol$_{(t-1)}$	0.212***	0.213***
	[0.029]	[0.029]
Smoke cigarette$_{(t-1)}$	0.002	0
	[0.055]	[0.056]
Real family income	0.002	0.002
	[0.005]	[0.005]
Peers use marijuana#Age		
Non of them#Age		*Reference*
A few of them#Age		−0.003
		[0.004]
Most of them#Age		−0.006
		[0.005]
All of them#Age		0.03
		[0.021]
Constant	−0.111	−0.146
	[0.065]	[0.076]
N	900	900
AIC	−4100	−4100
BIC	−4000	−4000
sigma_u	0.023	0.022
sigma_e	0.028	0.028
rho	0.388	0.384
R-squared Overall	0.877	0.878

$^{*}p < 0.05, ^{**}p < 0.01, ^{***}p < 0.001$

The standard deviation is inside the bracket [...]

Table B.14: Fixed effect linear model, NSDUH weighted pseudo-panel investigating the age, peers and last year marijuana use interaction

	Final model	extra information
Drug info$_{(t-1)}$=1	0.042	0.055
	[0.059]	[0.082]
Age	0.013***	0.020**
	[0.004]	[0.006]
Male # Drug info$_{(t-1)}$=1	−0.013	−0.01
	[0.010]	[0.010]
Non-Black, non-Hispanic # Drug info$_{(t-1)}$=1	*Reference*	*Reference*
Black# Drug info$_{(t-1)}$=1	−0.016	−0.013
	[0.013]	[0.013]
Hispanic# Drug info$_{(t-1)}$=1	−0.015	−0.007
	[0.013]	[0.013]
Age# Drug info$_{(t-1)}$=1	−0.001	−0.005
	[0.003]	[0.005]
The reported risk perception		
No risk	*Reference*	*Reference*
Slight risk	−0.043***	−0.082***
	[0.005]	[0.023]
Moderate risk	−0.069***	−0.097***
	[0.006]	[0.025]
Great risk	−0.076***	−0.095**
	[0.007]	[0.030]
Caring family$_{(t-1)}$	−0.022	−0.019
	[0.012]	[0.012]
Peers use marijuana		
None of them	*Reference*	*Reference*
A few of them	0.007*	0.049
	[0.003]	[0.053]
Most of them	0.038***	0.129
	[0.007]	[0.080]
All of them	0.092***	−0.436
	[0.022]	[0.374]
Marijuana use$_{(t-1)}$	0.184***	0.089
	[0.047]	[0.098]
Drink alcohol$_{(t-1)}$	0.212***	0.211***
	[0.029]	[0.029]
Smoke cigarette$_{(t-1)}$	0.002	0.005
	[0.055]	[0.056]
Real family income	0.002	0.003
	[0.005]	[0.005]
Live in metropolitan city=1	0	0
	[0.003]	[0.003]
Drug info$_{(t-1)}$=1 # Marijuana use$_{(t-1)}$		0.104
		[0.094]
Peers use marijuana#Age		
None of them# Age		*Reference*
A few of them# Age		−0.003
		[0.004]
Most of them#Age		−0.006
		[0.005]
All of them# Age		0.032
		[0.023]
The reported risk perception#Drug info$_{(t-1)}$=1		
No risk#Drug info$_{(t-1)}$=1		*Reference*
Slight risk#Drug info$_{(t-1)}$=1		0.041
		[0.023]
Moderate risk#Drug info$_{(t-1)}$=1		0.029
		[0.025]
Great risk# Drug info$_{(t-1)}$=1		0.02
		[0.030]
Constant	−0.111	−0.175
	[0.065]	[0.096]
N	900	900
AIC	−4100	−4100
BIC	−4000	−4100
sigma_u	0.023	0.022
sigma_e	0.028	0.028
rho	0.388	0.378
R-squared Overall	0.877	0.88

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$
The standard deviation is inside the bracket [...]

Table B.15: Fixed effect linear model, NSDUH weighted pseudo-panel investigating the all age levels, peers and last year marijuana use interaction

	extra information
Drug info$_{(t-1)}$=1	-0.006
	[0.029]
Age=13	*Reference*
Age=14	0.01
	[0.014]
Age=15	0.058^{**}
	[0.019]
Age=16	0.124^{**}
	[0.039]
Age=17	0.165^{***}
	[0.031]
Marijuana use$_{(t-1)}$	0.069
	[0.087]
The reported risk perception	
No risk	*Reference*
Slight risk	-0.077^{***}
	[0.022]
Moderate risk	-0.096^{***}
	[0.023]
Great risk	-0.095^{**}
	[0.029]
Male # Drug info$_{(t-1)}$=1	-0.008
	[0.010]
Non-Black, non-Hispanic # Drug info$_{(t-1)}$=1	*Reference*
Black# Drug info$_{(t-1)}$=1	-0.015
	[0.013]
Hispanic# Drug info$_{(t-1)}$=1	-0.01
	[0.013]
Age=13#Drug info$_{(t-1)}$=1	*Reference*
Age=14#Drug info$_{(t-1)}$=1	0.006
	[0.014]
Age=15#Drug info$_{(t-1)}$=1	0.004
	[0.018]
Age=16#Drug info$_{(t-1)}$=1	-0.024
	[0.020]
Age=17#Drug info$_{(t-1)}$=1	-0.017
	[0.020]
Marijuana use$_{(t-1)}$#Drug info$_{(t-1)}$=1	0.126
	[0.087]
Risk perception# Drug info$_{(t-1)}$	
No risk #Drug info$_{(t-1)}$=1	*Reference*
Slight risk #Drug info$_{(t-1)}$=1	0.037
	[0.023]
Moderate risk #Drug info$_{(t-1)}$=1	0.028
	[0.023]
Great risk #Drug info$_{(t-1)}$=1	0.023
	[0.028]
Caring family$_{(t-1)}$	-0.019
	[0.012]
Peers use marijuana	
None of them	*Reference*
A few of them	-0.002
	[0.007]
Most of them	-0.062^{*}
	[0.025]
All of them	0.018^{*}
	[0.009]
Peers use marijuana# Age	
None of them # all Ages	*Reference*
A few of them # Age=13	*Reference*
A few of them # Age=14	0.004
	[0.008]
A few of them #Age=15	-0.022^{*}
	[0.010]
A few of them # Age=16	-0.053
	[0.033]
A few of them # Age=17	-0.094^{***}
	[0.027]
Most of them #Age=13	*Reference*
Most of them # Age=14	0.133^{**}
	[0.045]
Most of them # Age=15	0.066^{*}
	[0.026]
Most of them # Age=16	0.034
	[0.038]
Drank alcohol$_{(t-1)}$	0.211^{***}
	[0.028]
Smoke cigarette$_{(t-1)}$	0.012
	[0.058]
Real family income	0.003
	[0.005]
Live in metropolitan city=1	0.001
	[0.003]
Constant	0.081^{*}
	[0.036]
N	900
AIC	-4100
BIC	-3900
sigma_u	0.023
sigma_e	0.028
rho	0.387
R-squared Overall	0.88

$^{*}p < 0.05, ^{**}p < 0.01, ^{***}p < 0.001$
The standard deviation is inside the bracket [...]

Table B.16: The reported risk perception and age interaction with school-based drug education (2002-2014)
fixed effect linear weighted NSDUH pseudo-panel data

	Final model with categorical age
Drug info$_{(t-1)}$=1	0.524*
	[0.204]
Age	
13	*Reference*
14	-0.599^{**}
	[0.201]
15	-0.941^{***}
	[0.230]
16	-1.075^{***}
	[0.203]
17	-1.469^{***}
	[0.201]
Drug info$_{(t-1)}$=1#13	*Reference*
Drug info$_{(t-1)}$=1#14	0.153
	[0.201]
Drug info$_{(t-1)}$=1#15	−0.177
	[0.216]
Drug info$_{(t-1)}$=1#16	-0.370^{*}
	[0.175]
Drug info$_{(t-1)}$=1#17	−0.33
	[0.168]
Male # Drug info$_{(t-1)}$=1	−0.036
	[0.123]
Non-Black, non-Hispanic # Drug info$_{(t-1)}$=1	*Reference*
Black# Drug info$_{(t-1)}$=1	-0.390^{*}
	[0.154]
Hispanic#Drug info$_{(t-1)}$=1	-0.471^{**}
	[0.176]
Caring family$_{(t-1)}$	0.499*
	[0.223]
Peers use marijuana	
None of them	*Reference*
A few of them	−0.005
	[0.089]
Most of them	0.043
	[0.122]
All of them	−0.305
	[0.188]
Real family income	−0.011
	[0.082]
Live in metropolitan city=1	−0.029
	[0.055]
Constant	2.178***
	[0.425]
N	900
AIC	918.53
BIC	1004.973
sigma_u	0.701
sigma_e	0.459
rho	0.7
R-squared	
Overall	0.387

$^{*}p < 0.05,^{**}p < 0.01,^{***}p < 0.001$
The standard deviation is inside the bracket [...]

Table B.17: The reported risk perception and peers interaction with school-based drug education (2002-2014)
fixed effect linear weighted NSDUH pseudo-panel data

	Final model with categorical age
Drug $\text{info}_{(t-1)}$=1	0.497^{*}
	[0.229]
Age	
13	*Reference*
14	-0.486^{*}
	[0.213]
15	-0.808^{**}
	[0.276]
16	-0.909^{***}
	[0.252]
17	-1.309^{***}
	[0.246]
Peers use marijuana	
None of them	*Reference*
A few of them	−0.161
	[0.203]
Most of them	−0.4
	[0.244]
All of them	−0.426
	[0.365]
Caring $\text{family}_{(t-1)}$	0.241
	[0.305]
Male # Drug $\text{info}_{(t-1)}$=1	−0.042
	[0.128]
Non-Black, non-Hispanic # Drug $\text{info}_{(t-1)}$=1	*Reference*
Black# Drug $\text{info}_{(t-1)}$=1	-0.523^{**}
	[0.169]
Hispanic#Drug $\text{info}_{(t-1)}$=1	-0.520^{**}
	[0.182]
Drug $\text{info}_{(t-1)}$=1#13	*Reference*
Drug $\text{info}_{(t-1)}$=1#14	0.039
	[0.224]
Drug $\text{info}_{(t-1)}$=1#15	−0.311
	[0.281]
Drug $\text{info}_{(t-1)}$=1#16	-0.532^{*}
	[0.252]
Drug $\text{info}_{(t-1)}$=1#17	-0.488^{*}
	[0.247]
Peers use marijuana	
None of them# Drug $\text{info}_{(t-1)}$=1	Reference
A few of them# Drug $\text{info}_{(t-1)}$=1	0.165
	[0.224]
Most of them# Drug $\text{info}_{(t-1)}$=1	0.469
	[0.261]
All of them# Drug $\text{info}_{(t-1)}$=1	0.04
	[0.413]
Caring $\text{family}_{(t-1)}$# Drug $\text{info}_{(t-1)}$=1	0.294
	[0.307]
Real family income	−0.009
	[0.083]
Live in metropolitan city=1	−0.03
	[0.055]
Constant	2.213^{***}
	[0.424]
N	900
AIC	923.92
BIC	1029.573
sigma_u	0.701
sigma_e	0.46
rho	0.699
R-squared	
Overall	0.391

$^{*}p < 0.05, ^{**}p < 0.01, ^{***}p < 0.001$
The standard deviation is inside the bracket [...]

Table B.18: Age-year marijuana use percentage compression between NSDUH pseudo-panel (2002-2014) and NLSY97

	NSDUH Pseudo-Panel												NLSY97
Age	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	
13	1.33%	1.84%	1.68%	1.79%	2.20%	1.81%	1.86%	2.16%	2.69%	2.07%	2.09%	2.12%	4.13%
14	6.26%	7.20%	5.54%	6.01%	5.39%	6.28%	8.50%	6.31%	8.37%	5.72%	7.34%		19.25%
15	12.74%	13.05%	10.86%	13.00%	14.85%	14.87%	15.08%	12.57%	14.95%	15.34%			26.67%
16	17.36%	19.03%	18.11%	18.43%	22.27%	21.57%	20.86%	22.93%	22.78%				31.34%
17	23.05%	23.79%	23.88%	25.84%	25.67%	27.18%	28.19%	27.42%					30.80%

Table B.19: Age heterogeneity and marijuana use
linear models, pooled NSDUH pseudo-panel (2002-2014)

Robust	Age 13		Age 14		Age 15		Age 16		Age 17	
Marijuana use$_{(t)}$	Males	Females	Males	Females	Males	Females	Males	Females	Males	Females
Drug education at school$_{(t-1)}$	−0.04	−0.012	0.054	−0.069	0.022	−0.03	0.102	0.076	−0.132	0.123
	[0.030]	[0.031]	[0.076]	[0.067]	[0.085]	[0.072]	[0.092]	[0.136]	[0.105]	[0.112]
The reported risk perception	−0.043	−0.048	−0.074	−0.134*	−0.207**	−0.243**	−0.193*	−0.316***	−0.244*	−0.513***
	[0.033]	[0.036]	[0.048]	[0.067]	[0.073]	[0.080]	[0.077]	[0.090]	[0.097]	[0.092]
Caring family$_{(t-1)}$	−0.017	0.01	0.025	0.024	−0.076	0.051	0.003	−0.016	−0.063	−0.073
	[0.015]	[0.016]	[0.023]	[0.038]	[0.041]	[0.030]	[0.042]	[0.050]	[0.037]	[0.039]
Peers use marijuana	−0.002	0.194**	0.079	0.322*	0.136	0.269*	0.542**	0.316	0.336*	0.540**
	[0.050]	[0.065]	[0.086]	[0.126]	[0.144]	[0.125]	[0.183]	[0.176]	[0.164]	[0.193]
Marijuana use$_{(t-1)}$	0	0	0.29	−0.205	0.127	0.145	−0.084	0.403**	−0.05	0.127
	[.]	[.]	[0.181]	[0.200]	[0.182]	[0.132]	[0.121]	[0.151]	[0.143]	[0.105]
drank alcohol$_{(t-1)}$	−0.032	0.086	0.092	0.144*	0.072	0.155	0.079	0.078	−0.109	0.05
	[0.067]	[0.051]	[0.071]	[0.066]	[0.089]	[0.089]	[0.090]	[0.112]	[0.094]	[0.101]
Smoked cigarettes$_{(t-1)}$	0.062	−0.119	0.154	−0.094	−0.051	0.127	0.185	−0.232	0.372*	0.048
	[0.089]	[0.085]	[0.118]	[0.133]	[0.146]	[0.124]	[0.146]	[0.138]	[0.145]	[0.125]
Like to do risky things										
Never	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*	*Reference*
Seldom	0.011	−0.004	0.01	0.012	0	0.023	0.059**	−0.003	0.064**	−0.021
	[0.006]	[0.007]	[0.012]	[0.017]	[0.019]	[0.021]	[0.021]	[0.029]	[0.023]	[0.032]
Sometimes / always	0.019*	0.013	0.031	0.05	0.014	0.079*	0.077**	0.079	0.093**	0.005
	[0.008]	[0.010]	[0.016]	[0.026]	[0.029]	[0.034]	[0.027]	[0.048]	[0.031]	[0.048]
Real family Income	−0.006	0.006	−0.003	0.005	0.004	0.008	−0.047**	0.029	0.014	0.041*
	[0.005]	[0.006]	[0.009]	[0.011]	[0.014]	[0.017]	[0.015]	[0.018]	[0.018]	[0.020]
Live in metropolitan city	0.097*	0.049	−0.08	0.132	−0.122	−0.124	0.172	−0.029	0.073	−0.366*
	[0.039]	[0.043]	[0.066]	[0.085]	[0.133]	[0.112]	[0.143]	[0.140]	[0.117]	[0.145]
Black	−0.029*	−0.005	0.011	−0.02	0.028	0.018	−0.112**	0.022	0.021	0.106*
	[0.013]	[0.016]	[0.022]	[0.029]	[0.040]	[0.042]	[0.042]	[0.048]	[0.042]	[0.050]
Hispanic	−0.026*	0.002	0.017	−0.022	0.018	0.032	−0.116**	0.036	0.01	0.091
	[0.012]	[0.015]	[0.020]	[0.027]	[0.038]	[0.038]	[0.039]	[0.046]	[0.039]	[0.046]
Constant	0.029	−0.05	0.074	−0.044	0.228	0.16	0.022	−0.017	0.111	0.172
	[0.039]	[0.046]	[0.093]	[0.107]	[0.141]	[0.123]	[0.142]	[0.200]	[0.146]	[0.188]
N	108	108	99	99	90	90	81	81	72	72
R-squared	0.373	0.564	0.597	0.726	0.653	0.81	0.82	0.845	0.885	0.827

$^{*} p < 0.05, ^{**} p < 0.01, ^{***} p < 0.001$
The standard deviation is inside the bracket [...]
IN this table I did not use weighted categorized NSDUH pseudo panel.

Table B.20: Effect of school-based drug education on marijuana use in the coming year random effect linear model, NSDUH pseudo-panel data (2002-2014)

Robust	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6		Model 7	
Marijuana use$_{(t)}$	Males	Females	Males	Females	Males	Females	Males	Females	Males	Females	Males	Females	Male	Female
Drug education at school$_{(t-1)}$	-0.230^{**}	-0.462^{***}	-0.074	-0.174^{**}	-0.089^{*}	-0.110^{*}	-0.071^{*}	-0.087	0.036	0.045	0.012	0.007	0.011	-0.046
	[0.083]	[0.076]	[0.042]	[0.058]	[0.036]	[0.054]	[0.033]	[0.046]	[0.028]	[0.051]	[0.028]	[0.050]	[0.040]	[0.052]
The reported risk perception			-0.873^{***}	-0.991^{***}	-0.573^{***}	-0.572^{***}	-0.318^{***}	-0.430^{***}	-0.164^{***}	-0.265^{***}	-0.187^{***}	-0.308^{***}	-0.185^{***}	-0.295^{***}
			[0.026]	[0.043]	[0.043]	[0.056]	[0.039]	[0.055]	[0.026]	[0.033]	[0.027]	[0.032]	[0.030]	[0.035]
Caring family$_{(t-1)}$					-0.171^{***}	-0.181^{***}	-0.113^{***}	-0.126^{***}	-0.050^{***}	-0.049^{***}	-0.014	-0.008	-0.03	-0.009
					[0.016]	[0.017]	[0.017]	[0.019]	[0.012]	[0.011]	[0.013]	[0.011]	[0.019]	[0.017]
Peers use marijuana							0.354^{***}	0.253^{***}	0.243^{***}	0.164^{***}	0.219^{***}	0.113^{***}	0.137^{**}	0.246^{***}
							[0.035]	[0.042]	[0.026]	[0.030]	[0.026]	[0.030]	[0.048]	[0.047]
Marijuana use$_{(t-1)}$									0.583^{***}	0.607^{***}	0.401^{***}	0.341^{***}	0.350^{***}	0.392^{***}
									[0.042]	[0.054]	[0.049]	[0.082]	[0.055]	[0.078]
drank alcohol$_{(t-1)}$											0.171^{***}	0.227^{***}	0.153^{***}	0.243^{***}
											[0.033]	[0.048]	[0.035]	[0.046]
Smoked cigarettes$_{(t-1)}$											0.008	0.006	-0.035	-0.094
											[0.068]	[0.076]	[0.077]	[0.078]
Like to do risky things														
Never													*Reference*	*Reference*
Seldom													0.004	-0.019^{*}
													[0.006]	[0.008]
Sometimes / always													0.011	-0.006
													[0.009]	[0.012]
Real family Income													-0.002	0.018^{**}
													[0.006]	[0.006]
Live in metropolitan city													0.02	0.002
													[0.049]	[0.056]
Age													0.011^{*}	-0.012^{**}
													[0.005]	[0.004]
Black													-0.007	0.018
													[0.016]	[0.016]
Hispanic													-0.015	0.011
													[0.016]	[0.014]
Constant	0.297^{***}	0.492^{***}	0.670^{***}	0.866^{***}	0.524^{***}	0.562^{***}	0.244^{***}	0.360^{***}	0.068^{*}	0.144^{**}	0.087^{**}	0.189^{***}	-0.043	0.274^{**}
	[0.066]	[0.063]	[0.033]	[0.048]	[0.030]	[0.046]	[0.036]	[0.049]	[0.029]	[0.044]	[0.029]	[0.045]	[0.072]	[0.084]
N	450	450	450	450	450	450	450	450	450	450	450	450	450	450
sigma_u	0.011	0.049	0.039	0.052	0.027	0.032	0.018	0.026	0	0	0	0	0	0
sigma_e	0.1	0.097	0.045	0.051	0.041	0.046	0.04	0.046	0.036	0.041	0.035	0.039	0.035	0.039
rho	0.011	0.204	0.435	0.505	0.305	0.331	0.163	0.25	0	0	0	0	0	0
R-squared														
Overall	0.014	0.032	0.667	0.603	0.768	0.751	0.824	0.778	0.885	0.871	0.895	0.89	0.899	0.9

$^{*}p < 0.05, ^{**}p < 0.01, ^{***}p < 0.001$
The standard deviation is inside the bracket [...]
Hausman test leads to reject the null hypothesis that Random effect provides consistent estimates.As result, in chapter 1 fixed effect linear model is analyzed.
IN this table I did not use weighted categorized NSDUH pseudo panel.

Table B.21: Effect of school-based drug education on the reported risk perception random effect linear model, NSDUH pseudo-panel data (2002-2014)

Risk Perception of Monthly marijuana use	Model 1 Males	Model 1 Females	Model 2 Males	Model 2 Females	Model 3 Males	Model 3 Females	Model 4 Males	Model 4 Females	Model 5 Males	Model 5 Females
Drug education at school$_{(t-1)}$	0.410^{***}	0.390^{***}	0.261^{***}	0.142^{**}	0.154^{**}	0.086	0.226^{***}	0.149^{**}	0.173^{**}	0.140^{**}
	[0.095]	[0.062]	[0.070]	[0.047]	[0.053]	[0.045]	[0.057]	[0.049]	[0.056]	[0.054]
Caring family$_{(t-1)}$			0.334^{***}	0.233^{***}	0.110^{***}	0.093^{***}	0.083^{***}	0.057^{**}	0.057^{*}	0.041
			[0.017]	[0.014]	[0.019]	[0.018]	[0.020]	[0.019]	[0.022]	[0.023]
Peers use marijuana					-0.559^{***}	-0.393^{***}	-0.573^{***}	-0.424^{***}	-0.321^{***}	-0.319^{***}
					[0.041]	[0.038]	[0.041]	[0.038]	[0.067]	[0.056]
Like to do risky things										
Never							*Reference*	*Reference*	*Reference*	*Reference*
Seldom							-0.039^{***}	-0.033^{**}	-0.083^{***}	-0.077^{***}
							[0.011]	[0.012]	[0.012]	[0.011]
Sometimes / always							-0.045^{**}	-0.050^{***}	-0.101^{***}	-0.091^{***}
							[0.014]	[0.015]	[0.014]	[0.016]
Real family Income									0.045^{***}	0.060^{***}
									[0.009]	[0.007]
Live in metropolitan city									-0.134	-0.135
									[0.072]	[0.083]
Age									-0.024^{***}	-0.008
									[0.006]	[0.005]
Black									0.100^{***}	0.148^{***}
									[0.028]	[0.020]
Hispanic									0.105^{***}	0.136^{***}
									[0.025]	[0.018]
Constant	0.259^{***}	0.304^{***}	0.348^{***}	0.499^{***}	0.629^{***}	0.693^{***}	0.611^{***}	0.684^{***}	0.826^{***}	0.586^{***}
	[0.075]	[0.050]	[0.054]	[0.038]	[0.046]	[0.039]	[0.048]	[0.041]	[0.111]	[0.124]
N	450	450	450	450	450	450	450	450	450	450
sigma_u	0.045	0.052	0.048	0.047	0.05	0.049	0.035	0.035	0.034	0.019
sigma_e	0.096	0.08	0.061	0.052	0.046	0.045	0.046	0.045	0.043	0.043
rho	0.18	0.298	0.38	0.444	0.537	0.546	0.366	0.375	0.377	0.165
R-squared										
Overall	0.02	0.052	0.495	0.491	0.654	0.561	0.684	0.6	0.725	0.668

$^{*} p < 0.05, ^{**} p < 0.01, ^{***} p < 0.001$
The standard deviation is inside the bracket [...]
Hausman test leads to reject the null hypothesis that Random effect provides consistent estimates.As result, in chapter 1 fixed effect linear model is analyzed.

Table B.22: Grade levels in different countries

Starting age	Canada	United States	Britain	France	Germany	Poland
0+	Nursery	Nursery	Nursery		Kinderkrippe	Żłobek
3-4	Preschool	Preschool	Nursery	Petite section	Kindergarten	Przedszkole
4-5	Junior kindergarten	Preschool	Reception	Moyenne section	Kindergarten	Przedszkole
5-6	Senior kindergarten	Kindergarten	Year 1	Grande section	Kindergarten	Zerówka
6-7	Grade 1	Grade 1	Year 2	Cours préparatoire	Grundschule	Klasa 1
7-8	Grade 2	Grade 2	Year 3	Cours élémentaire première année	Grundschule	Klasa 2
8-9	Grade 3	Grade 3	Year 4	Cours élémentaire deuxième année	Grundschule	Klasa 3
9-10	Grade 4	Grade 4	Year 5	Cours moyen première année	Grundschule	Klasa 4
10-11	Grade 5	Grade 5	Year 6	Course moyen deuxième année	Hauptschule	Klasa 5
11-12	Grade 6	Grade 6	Year 7	Sixième	Hauptschule	Klasa 6
12-13	Grade 7	Grade 7	Year 8	Cinquième	Hauptschule	Klasa 7
13-14	Grade 8	Grade 8	Year 9	Quatrième	Hauptschule	Klasa 8
14-15	Grade 9	Grade 9	Year 10	Troisième	Hauptschule	Liceum 1
15-16	Grade 10	Grade 10	Year 11	Seconde	Hauptschule	Liceum 2
16-17	Grade 11	Grade 11	Year 12	Première	Höhere Handelsschule	Liceum 3
17-18	Grade 12	Grade 12	Year 13	Terminale		Liceum 4

Reference (https://www.ourkids.net/school/Canada-grade-levels)

Appendix C

Information Regarding Chapter 2

Table C.1: Parameter estimates and marginal effects for Model 1 in Table 2.4

	Grade transitions			
Variable	Estimate	Std err	Marginal effect	Std err
Male	-0.2089	0.0309	-0.032	0.005
Black	0.2286	0.0492	0.037	0.008
Hispanic	0.0341	0.0512	0.008	0.009
Nuclear	0.3978	0.0346	0.063	0.006
Family income (in $10,000)	0.0263	0.0029	0.004	0.001
Mother high school graduate	0.3201	0.0524	0.049	0.009
Mother college graduate	0.7009	0.0525	0.111	0.009
Standardized AFQT scores	0.6064	0.0203	0.097	0.003
Mother's age at birth	0.0268	0.0033	0.004	0.001
Urban	-0.0291	0.0351	-0.005	0.005
Number of siblings	-0.0391	0.0151	-0.007	0.002
I(marijuana use period t)*I(highest grade completed=12)	0.7126	0.0644	0.114	0.011
I(marijuana use period t)*I(12<highest grade completed<16)	0.9698	0.0592	0.156	0.009
I(marijuana use period t)*I(15<highest grade completed)	1.2502	0.0876	0.199	0.015
I(highest grade completed=12)	-2.0696	0.0499	-0.331	0.008
I(12<highest grade completed<16)	-1.5107	0.047	-0.242	0.007
I(15<highest grade completed)	-3.2035	0.0687	-0.512	0.011
I(marijuana use period t)	-1.4763	0.0421	-0.237	0.006
Highest grade completed age 15	0.3728	0.0243	0.06	0.004
Constant	-3.9561	0.2711		

Table C.2: Parameter estimates and marginal effects for Model 2 in Table 2.4

	Grade transitions				Transitions into marijuana use			
Variable	Estimate	Std err	Marginal effect	Std err	Estimate	Std err	Marginal effect	Std err
Male	-0.263	0.0583	-0.028	0.007	0.0464	0.0797	0.004	0.005
Black	0.4474	0.0868	0.048	0.009	-0.7511	0.1316	-0.043	0.008
Hispanic	0.2015	0.0926	0.022	0.009	-0.1107	0.1354	-0.004	0.008
Nuclear	0.5182	0.0642	0.055	0.006	-0.602	0.0903	-0.035	0.006
Family income (in $10,000)	0.0384	0.0048	0.004	0.001	0.0115	0.0072	0.001	0.001
Mother high school graduate	0.6486	0.0961	0.067	0.011	-0.0034	0.1359	-0.002	0.008
Mother college graduate	0.9816	0.0914	0.104	0.01	0.151	0.1369	0.008	0.008
Standardized AFQT scores	0.9043	0.0373	0.096	0.004	-0.1875	0.0493	-0.011	0.003
Mother's age at birth	0.0443	0.0063	0.005	0.001	-0.0306	0.0087	-0.002	0.0005
Urban	-0.0295	0.0686	-0.002	0.007	0.3633	0.092	0.02	0.005
Number of siblings	-0.1097	0.0302	-0.012	0.003	-0.2022	0.04	-0.012	0.002
I(marijuana use period t)∗I(highest grade completed=12)	0.7058	0.112	0.075	0.012				
I(marijuana use period t)∗I(12<highest grade completed<16)	0.9539	0.1111	0.101	0.012				
I(marijuana use period t)∗I(15<highest grade completed)	1.2234	0.1265	0.13	0.014				
I(highest grade completed=12)	-2.9862	0.0894	-0.318	0.008				
I(12<highest grade completed<16)	-3.2309	0.0965	-0.343	0.009				
I(15<highest grade completed)	-6.3192	0.1193	-0.671	0.01				
I(marijuana use period t)	-1.4545	0.0854	-0.156	0.009				
Highest grade completed age 15	0.7541	0.0467	0.08	0.005				
Log duration					1.2038	0.8143		
Log duration squared					2.5215	0.5282		
Log duration cubed					-0.8304	0.1156		
First masspoint	-9.1869	0.5306			-8.6886	0.5627		
Second masspoint	-5.897	0.524			-5.6127	0.508		
P1	-0.1684	0.1766						
P2	0.0442	0.1814						
P3	0.3355	0.0831						

Table C.3: Parameter estimates and marginal effects for Model 3 in Table 2.4

	Grade transitions				Transitions into marijuana use			
Variable	Estimate	Std err	Marginal effect	Std err	Estimate	Std err	Marginal effect	Std err
Male	-0.2149	0.0535	-0.022	0.006	0.0464	0.0671	0.003	0.004
Black	0.2937	0.0876	0.032	0.009	-0.7518	0.1181	-0.043	0.007
Hispanic	0.1909	0.0807	0.018	0.007	-0.11	0.1131	-0.005	0.006
Nuclear	0.4149	0.0611	0.044	0.006	-0.5968	0.0847	-0.033	0.005
Family income (in $10,000)	0.0411	0.0047	0.004	0.001	0.0111	0.0071	0.001	0.001
Mother high school graduate	0.738	0.0666	0.079	0.005	-0.0076	0.0977	0.001	0.005
Mother college graduate	1.0966	0.0655	0.116	0.006	0.138	0.1042	0.009	0.005
Standardized AFQT scores	0.8655	0.0374	0.092	0.004	-0.1866	0.0444	-0.011	0.003
Mother's age at birth	0.0479	0.0062	0.005	0.001	-0.0309	0.0081	-0.002	0.001
Urban	-0.0905	0.0631	-0.01	0.006	0.3642	0.0822	0.022	0.005
Number of siblings	-0.0745	0.029	-0.008	0.003	-0.2086	0.0378	-0.012	0.002
I(marijuana use period t)∗I(highest grade completed=12)	0.5044	0.1035	0.055	0.011				
I(marijuana use period t)∗I(12<highest grade completed<16)	0.6786	0.1035	0.073	0.011				
I(marijuana use period t)∗I(15<highest grade completed)	0.7663	0.1231	0.081	0.013				
I(cigarette use period t)∗I(highest grade completed=12)	0.1623	0.0834	0.017	0.008				
I(cigarette use period t)∗I(12<highest grade completed<16)	0.2404	0.0939	0.025	0.01				
I(cigarette use period t)∗I(15<highest grade completed)	0.7636	0.1086	0.082	0.012				
I(alcohol use period t)∗I(highest grade completed=12)	0.7376	0.1782	0.079	0.019				
I(alcohol use period t)∗I(12<highest grade completed<16)	0.5207	0.1959	0.053	0.021				
I(alcohol use period t)∗I(15<highest grade completed)	0.6191	0.2725	0.064	0.03				
I(highest grade completed=12)	-3.6458	0.1593	-0.386	0.016				
I(12<highest grade completed<16)	-3.6515	0.1909	-0.383	0.019				
I(15<highest grade completed)	-7.1086	0.2568	-0.75	0.027				
I(marijuana use period t)	-0.9018	0.0879	-0.096	0.009				
I(cigarette use period t)	-0.9924	0.0783	-0.105	0.009				
I(alcohol use period t)	-0.8597	0.134	-0.091	0.014				
Highest grade completed age 15	0.7464	0.0458	0.078	0.004				
Log duration					1.2905	0.1421		
Log duration squared					2.4557	0.1263		
Log duration cubed					-0.8181	0.0518		
First masspoint	-8.0517	0.5376			-8.7984	0.2981		
Second masspoint	-4.8025	0.5294			-5.6311	0.3082		
P1	-0.1351	0.075						
P2	-0.076	0.0909						
P3	0.3572	0.0645						

Table C.4: Parameter estimates and marginal effects for Table 2.6

	Grade transitions				Transitions into marijuana use			
Variable	Estimate	Std err	Marginal effect	Std err	Estimate	Std err	Marginal effect	Std err
Male	-0.2051	0.0499	-0.019	0.005	0.049	0.0572	0.003	0.003
Black	0.1833	0.0756	0.018	0.007	-0.7355	0.101	-0.041	0.006
Hispanic	0.1264	0.0742	0.011	0.006	-0.1104	0.0874	-0.006	0.004
Nuclear	0.384	0.0601	0.036	0.005	-0.5879	0.0792	-0.032	0.005
Family income (in $10,000)	0.0432	0.0046	0.004	0.001	0.0109	0.0068	0.001	0.001
Mother high school graduate	0.7534	0.0808	0.072	0.006	-0.0084	0.0935	0	0.005
Mother college graduate	1.1671	0.0811	0.11	0.007	0.1254	0.0905	0.008	0.005
Standardized AFQT scores	0.8091	0.0369	0.076	0.003	-0.1769	0.0421	-0.01	0.003
Mother's age at birth	0.0474	0.0062	0.004	0.001	-0.0302	0.0071	-0.002	0.001
Urban	-0.059	0.0504	-0.006	0.005	0.3544	0.0755	0.021	0.004
Number of siblings	-0.0673	0.0317	-0.006	0.003	-0.2068	0.0342	-0.012	0.002
I(marijuana use period t)∗I(highest grade completed=12)∗I(age(m)<15)	-0.5281	0.118	-0.049	0.013				
I(marijuana use period t)∗I(12<highest grade completed<16)∗I(age(m)<15)	-0.491	0.1233	-0.046	0.012				
I(marijuana use period t)∗I(15<highest grade completed)∗I(age(m)<15)	-0.0582	0.1588	-0.005	0.017				
I(marijuana use period t)∗I(age(m)<15)	-0.3565	0.1001	-0.033	0.009				
I(marijuana use period t)∗I(highest grade completed=12)	0.5945	0.0834	0.055	0.007				
I(marijuana use period t)∗I(12<highest grade completed<16)	0.7398	0.0871	0.068	0.008				
I(marijuana use period t)∗I(15<highest grade completed)	0.7758	0.082	0.074	0.007				
I(cigarette use period t)∗I(highest grade completed=12)	0.1967	0.1034	0.017	0.01				
I(cigarette use period t)∗I(12<highest grade completed<16)	0.2601	0.1174	0.024	0.011				
I(cigarette use period t)∗I(15<highest grade completed)	0.7621	0.1205	0.072	0.012				
I(alcohol use period t)∗I(highest grade completed=12)	0.7166	0.1254	0.069	0.012				
I(alcohol use period t)∗I(12<highest grade completed<16)	0.494	0.1434	0.048	0.013				
I(alcohol use period t)∗I(15<highest grade completed)	0.6042	0.213	0.057	0.019				
I(highest grade completed=12)	-3.6164	0.1221	-0.339	0.012				
I(12<highest grade completed<16)	-3.6165	0.1412	-0.34	0.013				
I(15<highest grade completed)	-7.0987	0.209	-0.665	0.02				
I(marijuana use period t)	-0.8473	0.0723	-0.079	0.007				
I(cigarette use period t)	-0.9682	0.0959	-0.09	0.009				
I(alcohol use period t)	-0.8704	0.108	-0.084	0.01				
Highest grade completed age 15	0.7476	0.0411	0.07	0.003				
Log duration					1.3053	0.1383		
Log duration squared					2.4461	0.086		
Log duration cubed					-0.8258	0.0346		
First masspoint	-8.0789	0.5004			-9.087	0.3667		
Second masspoint	-4.8165	0.4923			-5.6835	0.2136		
P1	-0.3252	0.1395						
P2	-0.349	0.1538						
P3	0.231	0.0635						

Table C.5: Parameter estimates and marginal effects for males in Table 2.7

	Grade transitions				Transitions into marijuana use			
Variable	Estimate	Std err	Marginal effect	Std err	Estimate	Std err	Marginal effect	Std err
Black	0.0878	0.1355	0.01	0.014	-0.2786	0.1525	-0.016	0.009
Hispanic	0.2114	0.1239	0.023	0.012	0.0341	0.1534	0.003	0.008
Nuclear	0.5057	0.0974	0.053	0.01	-0.5755	0.112	-0.032	0.006
Family income (in $10,000)	0.0639	0.0071	0.006	0.001	0.022	0.0103	0.001	0.001
Mother high school graduate	0.6159	0.1246	0.064	0.012	-0.1421	0.1515	-0.008	0.008
Mother college graduate	0.9521	0.1283	0.099	0.013	-0.1399	0.1735	-0.008	0.009
Standardized AFQT scores	0.8578	0.0581	0.089	0.004	-0.1569	0.0601	-0.009	0.004
Mother's age at birth	0.0628	0.0092	0.006	0.001	-0.03	0.0113	-0.002	0.001
Urban	0.0296	0.1006	0.005	0.01	0.3315	0.1119	0.02	0.007
Number of siblings	-0.0097	0.0455	-0.002	0.005	-0.2243	0.0524	-0.012	0.003
I(marijuana use period t)∗I(highest grade completed=12)	0.7474	0.1522	0.075	0.014				
I(marijuana use period t)∗I(12<highest grade completed<16)	0.8735	0.1748	0.091	0.015				
I(marijuana use period t)∗I(15<highest grade completed)	1.0114	0.1859	0.105	0.021				
I(cigarette use period t)∗I(highest grade completed=12)	0.1597	0.1251	0.017	0.014				
I(cigarette use period t)∗I(12<highest grade completed<16)	0.1286	0.1556	0.014	0.015				
I(cigaretee use period t)∗I(15<highest grade completed)	1.0865	0.1723	0.112	0.018				
I(alcohol use period t)∗I(highest grade completed=12)	0.2652	0.279	0.03	0.026				
I(alcohol use period t)∗I(12<highest grade completed<16)	0.4472	0.2765	0.044	0.024				
I(alcohol use period t)∗I(15<highest grade completed)	-0.342	0.4023	-0.033	0.038				
I(highest grade completed=12)	-3.5834	0.258	-0.371	0.025				
I(12<highest grade completed<16)	-3.6951	0.2706	-0.38	0.024				
I(15<highest grade completed)	-6.5372	0.375	-0.673	0.035				
I(marijuana use period t)	-1.1266	0.129	-0.116	0.013				
I(cigarette use period t)	-1.291	0.1165	-0.133	0.013				
I(alcohol use period t)	-0.7465	0.1972	-0.076	0.018				
Highest grade completed age 15	0.5238	0.0666	0.054	0.006				
Log duration					0.1794	0.2576		
Log duration squared					3.1467	0.2506		
Log duration cubed					-0.9603	0.0839		
First masspoint	-6.2855	0.7674			-7.8407	0.3482		
Second masspoint	-3.0906	0.7608			-5.0079	0.3845		
P1	-0.0828	0.2795						
P2	-0.2816	0.3124						
P3	0.3804	0.1127						

Table C.6: Parameter estimates and marginal effects for females in Table 2.7

	Grade transitions				Transitions into marijuana use			
Variable	Estimate	Std err	Marginal effect	Std err	Estimate	Std err	Marginal effect	Std err
Black	0.3999	0.1109	0.042	0.011	-1.1333	0.1903	-0.063	0.012
Hispanic	0.1593	0.1231	0.018	0.012	-0.2519	0.1929	-0.013	0.01
Nuclear	0.4366	0.0807	0.047	0.008	-0.62	0.1347	-0.034	0.008
Family income (in $10,000)	0.0255	0.0063	0.003	0.001	0.0012	0.0096	0.001	0.001
Mother high school graduate	0.8659	0.1421	0.093	0.014	0.1237	0.1825	0.008	0.009
Mother college graduate	1.1494	0.1319	0.123	0.013	0.4353	0.1972	0.025	0.01
Standardized AFQT scores	0.8651	0.0525	0.092	0.004	-0.1672	0.0682	-0.009	0.004
Mother's age at birth	0.037	0.0084	0.004	0.001	-0.0259	0.0121	-0.001	0.001
Urban	-0.1371	0.0972	-0.012	0.01	0.4279	0.1283	0.025	0.008
Number of siblings	-0.0775	0.0397	-0.009	0.004	-0.1804	0.0532	-0.009	0.003
I(marijuana use period t)∗I(highest grade completed=12)	0.2895	0.1905	0.029	0.018				
I(marijuana use period t)∗I(12<highest grade completed<16)	0.5092	0.1749	0.055	0.015				
I(marijuana use period t)∗I(15<highest grade completed)	0.5614	0.1995	0.061	0.022				
I(cigarette use period t)∗I(highest grade completed=12)	0.1498	0.226	0.02	0.026				
I(cigarette use period t)∗I(12<highest grade completed<16)	0.2172	0.1913	0.024	0.022				
I(cigarette use period t)∗I(15<highest grade completed)	0.4836	0.2148	0.051	0.024				
I(alcohol use period t)∗I(highest grade completed=12)	1.1832	0.2865	0.124	0.024				
I(alcohol use period t)∗I(12<highest grade completed<16)	0.6327	0.305	0.063	0.03				
I(alcohol use period t)∗I(15<highest grade completed)	1.4691	0.4064	0.163	0.041				
I(highest grade completed=12)	-3.7295	0.2596	-0.395	0.024				
I(12<highest grade completed<16)	-3.6407	0.2894	-0.383	0.03				
I(15<highest grade completed)	-7.6789	0.4107	-0.819	0.039				
I(marijuana use period t)	-0.6931	0.1373	-0.074	0.015				
I(cigarette use period t)	-0.7986	0.1554	-0.086	0.019				
I(alcohol use period t)	-0.9175	0.2027	-0.098	0.02				
Highest grade completed age 15	0.8861	0.063	0.093	0.006				
Log duration					2.4237	1.7164		
Log duration squared					1.917	1.0029		
Log duration cubed					-0.7437	0.1945		
First masspoint	-9.5176	0.7071			-9.5937	1.0844		
Second masspoint	-6.1512	0.6961			-6.5976	1.0403		
P1	-0.1837	0.1672						
P2	0.1177	0.1788						
P3	0.3209	0.1239						

Table C.7: Parameter estimates and marginal effects for Table 2.8

	Grade transitions				Transitions into marijuana use			
Variable	Estimate	Std err	Marginal effect	Std err	Estimate	Std err	Marginal effect	Std err
Male	-0.2258	0.0478	-0.025	0.005	0.0449	0.0752	0.003	0.004
Black	0.3251	0.0894	0.035	0.01	-0.7508	0.1172	-0.043	0.007
Hispanic	0.2161	0.1	0.024	0.011	-0.1056	0.0982	-0.006	0.006
Nuclear	0.4546	0.0591	0.05	0.006	-0.599	0.0901	-0.035	0.005
Family income (in $10,000)	0.0399	0.0047	0.004	0.001	0.0113	0.0072	0.001	0.001
Mother high school graduate	0.7321	0.099	0.081	0.011	-0.0103	0.0828	-0.001	0.005
Mother college graduate	1.1047	0.0964	0.121	0.01	0.1398	0.0804	0.008	0.004
Standardized AFQT scores	0.8599	0.0366	0.094	0.004	-0.1869	0.0394	-0.011	0.002
Mother's age at birth	0.046	0.0062	0.005	0.001	-0.0309	0.0083	-0.002	0.001
Urban	-0.1134	0.0577	-0.012	0.007	0.3666	0.0897	0.021	0.005
Number of siblings	-0.0855	0.0273	-0.009	0.003	-0.2077	0.0364	-0.012	0.002
I(marijuana use period t)*I(highest grade completed=12)	0.6728	0.1163	0.014	0.046				
I(marijuana use period t)*I(12<highest grade completed<16)	0.7711	0.1069	0.085	0.012				
I(marijuana use period t)*I(highest grade completed=12)*I(Black)	-0.2122	0.2307	-0.023	0.025				
I(marijuana use period t)*I(12<highest grade completed<16)*I(Black)	-0.0318	0.2194	-0.002	0.024				
I(marijuana use period t)*I(highest grade completed=12)*I(Hispanic)	-1.0035	0.1379	-0.111	0.014				
I(marijuana use period t)*I(12<highest grade completed<16)*I(Hispanic)	-0.2514	0.1517	-0.028	0.017				
I(cigarette use period t)*I(highest grade completed=12)	0.1753	0.0844	0.019	0.009				
I(cigarette use period t)*I(12<highest grade completed<16)	0.4528	0.1094	0.05	0.012				
I(alcohol use period t)*I(highest grade completed=12)	0.7298	0.2082	0.079	0.023				
I(alcohol use period t)*I(12<highest grade completed<16)	0.4684	0.2445	0.05	0.027				
I(highest grade completed=12)	-3.6601	0.1824	-0.401	0.019				
I(12<highest grade completed<16)	-3.7787	0.2197	-0.414	0.023				
I(15<highest grade completed)	-2.9927	0.0724	-0.328	0.007				
I(marijuana use period t)	-0.9067	0.0881	-0.1	0.009				
I(marijuana use period t)*I(Black)	-0.0002	0.1832	0.001	0.02				
I(marijuana use period t)*I(Hispanic)	0.2466	0.1196	0.027	0.013				
I(cigarette use period t)	-0.9507	0.0901	-0.105	0.01				
I(alcohol use period t)	-0.8602	0.1718	-0.094	0.019				
Highest grade completed age 15	0.7475	0.0413	0.082	0.004				
Log duration					1.122	0.6822		
Log duration squared					2.5615	0.4413		
Log duration cubed					-0.8368	0.0981		
First masspoint	-8.0316	0.4815			-8.6571	0.5411		
Second masspoint	-4.7819	0.476			-5.5436	0.515		
P1	-0.0956	0.0461						
P2	0.0197	0.0468						
P3	0.4046	0.0613						

Table C.8: Parameter estimates and marginal effects for Table 2.9

	Grade transitions				Transitions into marijuana use			
Variable	Estimate	Std err	Marginal effect	Std err	Estimate	Std err	Marginal effect	Std err
Male	-0.2531	0.0562	-0.023	0.005	0.0499	0.0745	0.003	0.004
Black	0.2203	0.064	0.021	0.006	-0.7476	0.0971	-0.042	0.006
Hispanic	0.1453	0.0753	0.012	0.006	-0.1098	0.0891	-0.006	0.005
Nuclear	0.4317	0.0622	0.04	0.005	-0.5969	0.0863	-0.033	0.005
Family income (in $10,000)	0.0409	0.0047	0.004	0	0.0109	0.007	0.001	0
Mother high school graduate	0.6384	0.0609	0.06	0.005	-0.0039	0.0942	0	0.005
Mother college graduate	0.996	0.0616	0.092	0.006	0.1326	0.096	0.008	0.005
Standardized AFQT scores	0.8483	0.0359	0.079	0.003	-0.1826	0.0454	-0.01	0.003
Mother's age at birth	0.0437	0.006	0.004	0.001	-0.0306	0.0075	-0.002	0
Urban	-0.0103	0.0588	-0.001	0.005	0.3601	0.0811	0.021	0.004
Number of siblings	-0.1082	0.0262	-0.01	0.002	-0.2075	0.0368	-0.012	0.002
I(marijuana use period t)*I(highest grade completed=12)*I(poor)	-0.654	0.0876	-0.06	0.009				
I(marijuana use period t)*I(12<highest grade completed<16)*I(poor)	-0.5715	0.1284	-0.051	0.012				
I(marijuana use period t)*I(15<highest grade completed)*I(poor)	0.9079	0.1129	0.084	0.01				
I(marijuana use period t)*I(poor)	-0.2741	0.0882	-0.025	0.007				
I(marijuana use period t)*I(highest grade completed=12)	0.5616	0.091	0.053	0.008				
I(marijuana use period t)*I(12<highest grade completed<16)	0.6718	0.1112	0.062	0.01				
I(marijuana use period t)*I(15<highest grade completed)	0.66	0.1168	0.061	0.01				
I(cigarette use period t)*I(highest grade completed=12)	0.1555	0.1022	0.014	0.01				
I(cigarette use period t)*I(12<highest grade completed<16)	0.2397	0.1261	0.022	0.013				
I(cigarette use period t)*I(15<highest grade completed)	0.7856	0.1181	0.074	0.011				
I(alcohol use period t)*I(highest grade completed=12)	0.7114	0.1833	0.065	0.017				
I(alcohol use period t)*I(12<highest grade completed<16)	0.4863	0.2619	0.043	0.024				
I(alcohol use period t)*I(15<highest grade completed)	0.599	0.2325	0.056	0.022				
I(highest grade completed=12)	-3.5989	0.1549	-0.333	0.013				
I(12<highest grade completed<16)	-3.6107	0.2022	-0.333	0.018				
I(15<highest grade completed)	-7.0519	0.2157	-0.654	0.02				
I(marijuana use period t)	-0.8895	0.0844	-0.082	0.008				
I(cigarette use period t)	-0.9903	0.0866	-0.092	0.009				
I(alcohol use period t)	-0.8487	0.1392	-0.077	0.013				
Highest grade completed age 15	0.6948	0.042	0.064	0.003				
Log duration					1.2653	0.1259		
Log duration squared					2.4647	0.0827		
Log duration cubed					-0.8232	0.0387		
First masspoint	-7.3195	0.5064			-8.888	0.2766		
Second masspoint	-4.0589	0.503			-5.6315	0.2411		
P1	-0.2952	0.1053						
P2	-0.2394	0.0997						
P3	0.1953	0.0654						

Table C.9: Parameter estimates and marginal effects for Table 2.9, middle panel

	Grade transitions				Transitions into marijuana use			
Variable	Estimate	Std err	Marginal effect	Std err	Estimate	Std err	Marginal effect	Std err
Male	-0.2297	0.0545	-0.021	0.005	0.0477	0.0744	0.003	0.004
Black	0.2465	0.0838	0.025	0.008	-0.7525	0.0934	-0.043	0.005
Hispanic	0.1831	0.0764	0.015	0.006	-0.1114	0.1862	-0.005	0.01
Nuclear	0.4238	0.0713	0.04	0.006	-0.6005	0.0849	-0.033	0.005
Family income (in $10,000)	0.0416	0.0046	0.004	0.001	0.0111	0.0071	0.001	0.001
Mother high school graduate	0.7186	0.074	0.068	0.006	-0.008	0.1614	0.001	0.009
Mother college graduate	1.0687	0.0723	0.101	0.007	0.1348	0.125	0.009	0.007
Standardized AFQT scores	0.8617	0.0372	0.082	0.003	-0.1859	0.0464	-0.011	0.003
Mother's age at birth	0.0484	0.0063	0.004	0.001	-0.0313	0.0085	-0.002	0.001
Urban	-0.0725	0.0519	-0.008	0.005	0.3628	0.0848	0.022	0.005
Number of siblings	-0.0851	0.0543	-0.008	0.004	-0.2089	0.0382	-0.012	0.002
I(marijuana use period t)∗I(highest grade completed=12)∗I(single)	-0.2499	0.0825	-0.023	0.007				
I(marijuana use period t)∗I(12<highest grade completed<16)∗I(single)	-0.0852	0.0837	-0.008	0.008				
I(marijuana use period t)∗I(15<highest grade completed)∗I(single)	1.0481	0.288	0.095	0.025				
I(marijuana use period t)∗I(single)	-0.1624	0.0919	-0.017	0.008				
I(marijuana use period t)∗I(highest grade completed=12)	0.5418	0.1931	0.048	0.017				
I(marijuana use period t)∗I(12<highest grade completed<16)	0.6601	0.1585	0.06	0.014				
I(marijuana use period t)∗I(15<highest grade completed)	0.4665	0.0989	0.043	0.009				
I(cigarette use period t)∗I(highest grade completed=12)	0.1614	0.0774	0.015	0.008				
I(cigarette use period t)∗I(12<highest grade completed<16)	0.2271	0.1391	0.019	0.013				
I(cigarette use period t)∗I(15<highest grade completed)	0.7972	0.1066	0.075	0.011				
I(alcohol use period t)∗I(highest grade completed=12)	0.7449	0.1971	0.068	0.017				
I(alcohol use period t)∗I(12<highest grade completed<16)	0.541	0.1277	0.051	0.011				
I(alcohol use period t)∗I(15<highest grade completed)	0.6222	0.5194	0.068	0.048				
I(highest grade completed=12)	-3.6357	0.3162	-0.337	0.028				
I(12<highest grade completed<16)	-3.651	0.2441	-0.341	0.022				
I(15<highest grade completed)	-7.1005	0.3576	-0.674	0.034				
I(marijuana use period t)	-0.8347	0.1063	-0.076	0.01				
I(cigarette use period t)	-0.9868	0.0903	-0.092	0.008				
I(alcohol use period t)	-0.8741	0.1185	-0.083	0.011				
Highest grade completed age 15	0.7021	0.0641	0.065	0.006				
Log duration					1.2847	0.1675		
Log duration squared					2.4488	0.0921		
Log duration cubed					-0.8154	0.0405		
First masspoint	-7.5693	0.851			-8.7898	0.4246		
Second masspoint	-4.3143	0.8408			-5.6038	0.2551		
P1	-0.1703	0.2609						
P2	-0.1104	0.3326						
P3	0.3089	0.1405						

Table C.10: Parameter estimates and marginal effects for Table 2.9, bottom panel

	Grade transitions				Transitions into marijuana use			
Variable	Estimate	Std err	Marginal effect	Std err	Estimate	Std err	Marginal effect	Std err
Male	-0.2272	0.0482	-0.021	0.004	0.0473	0.055	0.003	0.003
Black	0.2689	0.0827	0.026	0.008	-0.751	0.0967	-0.043	0.006
Hispanic	0.169	0.0803	0.015	0.007	-0.1094	0.0843	-0.006	0.004
Nuclear	0.4301	0.0614	0.041	0.005	-0.5986	0.081	-0.033	0.005
Family income (in $10,000)	0.0418	0.0045	0.004	0.001	0.0111	0.0067	0.001	0.001
Mother high school graduate	0.6982	0.0925	0.068	0.008	-0.0059	0.0912	0.001	0.005
Mother college graduate	1.0308	0.0945	0.099	0.008	0.138	0.1229	0.008	0.007
Standardized AFQT scores	0.8693	0.0364	0.083	0.003	-0.1859	0.0403	-0.011	0.002
Mother's age at birth	0.0426	0.0062	0.004	0.001	-0.0309	0.0073	-0.002	0.001
Urban	-0.0615	0.0561	-0.006	0.005	0.3635	0.067	0.022	0.004
Number of siblings	-0.079	0.0273	-0.007	0.002	-0.2085	0.0333	-0.012	0.002
I(marijuana use period t)∗I(highest grade completed=12)∗I(teen)	-0.9473	0.1373	-0.087	0.014				
I(marijuana use period t)∗I(12<highest grade completed<16)∗I(teen)	0.0443	0.0784	0.004	0.007				
I(marijuana use period t)∗I(15<highest grade completed)∗I(teen)	1.0646	0.0637	0.101	0.006				
I(marijuana use period t)∗I(single)	-0.3948	0.1013	-0.036	0.009				
I(marijuana use period t)∗I(highest grade completed=12)	0.5554	0.0885	0.053	0.007				
I(marijuana use period t)∗I(12<highest grade completed<16)	0.6475	0.1057	0.061	0.009				
I(marijuana use period t)∗I(15<highest grade completed)	0.7169	0.1103	0.069	0.01				
I(cigarette use period t)∗I(highest grade completed=12)	0.1598	0.0499	0.014	0.005				
I(cigarette use period t)∗I(12<highest grade completed<16)	0.2075	0.0623	0.019	0.007				
I(cigarette use period t)∗I(15<highest grade completed)	0.7497	0.091	0.071	0.009				
I(alcohol use period t)∗I(highest grade completed=12)	0.7228	0.1423	0.069	0.012				
I(alcohol use period t)∗I(12<highest grade completed<16)	0.52	0.1411	0.049	0.013				
I(alcohol use period t)∗I(15<highest grade completed)	0.6055	0.2716	0.057	0.024				
I(highest grade completed=12)	-3.6167	0.1434	-0.343	0.012				
I(12<highest grade completed<16)	-3.6328	0.1517	-0.344	0.013				
I(15<highest grade completed)	-7.0824	0.2669	-0.671	0.025				
I(marijuana use period t)	-0.8706	0.0841	-0.083	0.008				
I(cigarette use period t)	-0.9822	0.0708	-0.093	0.007				
I(alcohol use period t)	-0.8435	0.0938	-0.08	0.008				
Highest grade completed age 15	0.7125	0.0414	0.067	0.004				
Log duration					1.2732	0.1182		
Log duration squared					2.4663	0.1272		
Log duration cubed					-0.8205	0.0504		
First masspoint	-7.5607	0.5078			-8.7991	0.2675		
Second masspoint	-4.2899	0.5013			-5.6242	0.3052		
P1	-0.1611	0.0648						
P2	-0.1023	0.1037						
P3	0.3229	0.0516						

Table C.11: Parameter estimates and marginal effects for Model in column 2, Table 2.10

	Grade transitions				Transitions into marijuana use			
Variable	Estimate	Std err	Marginal effect	Std err	Estimate	Std err	Marginal effect	Std err
Male	-0.2633	0.0578	-0.027	0.006	0.1642	0.0836	0.01	0.005
Black	0.3698	0.0934	0.039	0.01	-0.8137	0.1098	-0.046	0.006
Hispanic	0.2154	0.0881	0.035	0.013	-0.0346	0.1184	-0.002	0.007
Nuclear	0.3954	0.0622	0.042	0.006	-0.4726	0.0863	-0.027	0.005
Family income (in $10,000)	0.0448	0.0048	0.005	0.001	0.0143	0.0069	0.001	0.001
Mother high school graduate	0.7884	0.0624	0.082	0.006	-0.0098	0.1129	-0.001	0.006
Mother college graduate	1.0821	0.0637	0.113	0.006	0.2179	0.1127	0.013	0.006
Standardized AFQT scores	0.8378	0.037	0.087	0.004	-0.0536	0.0457	-0.003	0.003
Mother's age at birth	0.0422	0.0064	0.004	0.001	-0.02	0.0079	-0.001	0.001
Urban	-0.0645	0.0819	-0.007	0.009	0.2581	0.0882	0.014	0.005
Number of siblings	-0.1123	0.0325	-0.012	0.003	-0.1375	0.0369	-0.008	0.002
Peers	0.2629	0.0316	0.027	0.003	-0.5068	0.0446	-0.029	0.003
I(marijuana use period t)∗I(highest grade completed=12)	0.5027	0.1219	0.053	0.013				
I(marijuana use period t)∗I(12<highest grade completed<16)	0.7316	0.129	0.076	0.014				
I(marijuana use period t)∗I(15<highest grade completed)	0.7833	0.149	0.082	0.015				
I(cigarette use period t)∗I(highest grade completed=12)	0.1571	0.1186	0.016	0.012				
I(cigarette use period t)∗I(12<highest grade completed<16)	0.1848	0.1156	0.019	0.012				
I(cigarette use period t)∗I(15<highest grade completed)	0.8373	0.147	0.087	0.016				
I(alcohol use period t)∗I(highest grade completed=12)	0.6816	0.1494	0.071	0.016				
I(alcohol use period t)∗I(12<highest grade completed<16)	0.4251	0.1626	0.044	0.017				
I(alcohol use period t)∗I(15<highest grade completed)	0.6021	0.2744	0.062	0.029				
I(highest grade completed=12)	-3.629	0.1707	-0.378	0.018				
I(12<highest grade completed<16)	-3.6678	0.1671	-0.382	0.017				
I(15<highest grade completed)	-7.2324	0.2617	-0.754	0.026				
I(marijuana use period t)	-0.8965	0.0985	-0.094	0.01				
I(cigarette use period t)	-0.9894	0.1009	-0.103	0.011				
I(alcohol use period t)	-0.891	0.1227	-0.093	0.012				
Highest grade completed age 15	0.7546	0.0476	0.079	0.005				
Log duration					0.9135	0.1217		
Log duration squared					2.8312	0.1153		
Log duration cubed					-0.913	0.0435		
First masspoint	-7.9111	0.5696			-9.8094	0.399		
Second masspoint	-4.6286	0.5631			-6.2997	0.3098		
P1	-0.4767	0.1408						
P2	-0.3825	0.1354						
P3	0.2509	0.0622						

Table C.12: Parameter estimates and marginal effects for Model in column 1, Table 2.14

	Grade transitions				Transitions into marijuana use			
Variable	Estimate	Std err	Marginal effect	Std err	Estimate	Std err	Marginal effect	Std err
Male	-0.1981	0.0916	-0.02	0.009	0.207	0.1342	0.018	0.011
Black	0.5184	0.1333	0.053	0.013	-0.6514	0.2188	-0.053	0.017
Hispanic	0.6462	0.1376	0.065	0.013	0.2889	0.2429	0.024	0.02
Nuclear	0.5002	0.1046	0.051	0.011	-0.5852	0.1588	-0.048	0.014
Family income (in $10,000)	0.0547	0.0113	0.006	0.001	-0.0069	0.0134	-0.001	0.001
Mother high school graduate	0.6199	0.149	0.062	0.015	0.0915	0.2608	0.008	0.02
Mother college graduate	1.3056	0.1597	0.132	0.016	-0.0142	0.2787	-0.001	0.022
Standardized AFQT scores	0.6669	0.0603	0.067	0.006	-0.0204	0.1101	-0.002	0.009
Mother's age at birth	0.0386	0.0095	0.004	0.001	-0.0225	0.0139	-0.002	0.001
Urban	-0.0833	0.1199	-0.009	0.012	0.1965	0.1315	0.015	0.011
Number of siblings	-0.2214	0.0485	-0.022	0.005	-0.2637	0.0684	-0.022	0.005
Peers	0.3543	0.0496	0.036	0.005	-0.5391	0.077	-0.044	0.006
I(marijuana use period t)∗I(highest grade completed=12)	0.1253	0.1628	0.013	0.017				
I(marijuana use period t)∗I(12<highest grade completed<16)	0.5545	0.1497	0.056	0.016				
I(marijuana use period t)∗I(15<highest grade completed)	1.033	0.1747	0.104	0.017				
I(alcohol use period t)∗I(highest grade completed=12)	0.5796	0.2354	0.059	0.023				
I(alcohol use period t)∗I(12<highest grade completed<16)	0.0623	0.2225	0.006	0.022				
I(alcohol use period t)∗I(15<highest grade completed)	0.699	0.538	0.072	0.052				
I(highest grade completed=12)	-3.2621	0.21	-0.33	0.02				
I(12<highest grade completed<16)	-3.1278	0.2131	-0.316	0.02				
I(15<highest grade completed)	-6.8858	0.5195	-0.697	0.05				
I(marijuana use period t)	-1.0381	0.114	-0.105	0.011				
I(alcohol use period t)	-1.4809	0.175	-0.149	0.017				
Highest grade completed age 15	0.914	0.0714	0.093	0.007				
Log duration					0.1273	0.1794		
Log duration squared					1.5449	0.2067		
Log duration cubed					-0.5779	0.071		
First masspoint	-6.3671	0.8052			-4.6444	0.464		
Second masspoint	-9.7321	0.8211			-0.8348	0.4276		
P1	0.2358	0.1214						
P2	0.1758	0.1099						
P3	0.1402	0.0971						

Table C.13: Parameter estimates and marginal effects for Model in column 2, Table 2.14

	Grade transitions				Transitions into marijuana use			
Variable	Estimate	Std err	Marginal effect	Std err	Estimate	Std err	Marginal effect	Std err
Male	0.0244	0.105	0.002	0.01	0.8018	0.1964	0.029	0.007
Black	0.7171	0.1496	0.071	0.015	-0.7491	0.2366	-0.026	0.008
Hispanic	0.4859	0.1482	0.048	0.015	0.0398	0.1692	0.002	0.006
Nuclear	0.5228	0.1103	0.052	0.011	-0.4493	0.1059	-0.016	0.004
Family income (in $10,000)	0.08	0.0171	0.008	0.002	-0.0522	0.02	-0.002	0.001
Mother high school graduate	0.674	0.1461	0.067	0.014	-0.1497	0.0955	-0.005	0.003
Mother college graduate	1.3676	0.1468	0.136	0.014	-0.0763	0.0902	-0.003	0.003
Standardized AFQT scores	0.5906	0.0613	0.059	0.006	0.0847	0.0889	0.003	0.003
Mother's age at birth	0.0255	0.0129	0.003	0.001	-0.0082	0.0161	0	0.001
Urban	-0.0324	0.0396	-0.003	0.004	0.5423	0.192	0.019	0.007
Number of siblings	-0.0969	0.0566	-0.009	0.006	-0.2458	0.1276	-0.009	0.005
Peers	0.2693	0.0542	0.027	0.005	-0.825	0.125	-0.029	0.005
I(marijuana use period t)∗I(highest grade completed=12)	-0.0907	0.1652	-0.009	0.016				
I(marijuana use period t)∗I(12<highest grade completed<16)	-0.0678	0.1351	-0.007	0.013				
I(marijuana use period t)∗I(15<highest grade completed)	0.9897	0.1766	0.099	0.017				
I(alcohol use period t)∗I(highest grade completed=12)	0.7353	0.1704	0.073	0.016				
I(alcohol use period t)∗I(12<highest grade completed<16)	0.8242	0.1549	0.082	0.015				
I(alcohol use period t)∗I(15<highest grade completed)	0.7954	0.1777	0.079	0.018				
I(highest grade completed=12)	-2.921	0.1252	-0.29	0.012				
I(12<highest grade completed<16)	-3.0886	0.1468	-0.306	0.013				
I(15<highest grade completed)	-6.2514	0.1886	-0.621	0.013				
I(marijuana use period t)	-0.9025	0.1195	-0.09	0.012				
I(alcohol use period t)	-1.5034	0.1176	-0.149	0.012				
Highest grade completed age 15	0.9426	0.087	0.094	0.008				
Log duration					-0.4889	0.296		
Log duration squared					2.8152	0.3348		
Log duration cubed					-0.9605	0.1057		
First masspoint	-7.8855	0.9943			-7.1018	0.8056		
Second masspoint	-11.3219	1.0162			-2.6826	0.5994		
P1	1.1708	0.1319						
P2	1.1951	0.1206						
P3	0.1671	0.1242						

Appendix D

Information Regarding Chapter 3

D.1 Generation of Truncated Random Variables for the Simulated Likelihood Function

D.1.1 Binary Outcomes

In order to derive the likelihood function in equation (4), we need to generate random variables ($e_{i,t}$) from truncated standard normal distributions on $[L_{i,t}, U_{i,t}]$. This can be done by transformations of uniformly distributed random variables, $u_{i,t} \sim U[0,1]$. Specifically, for each independent simulation run $(j), e_{i,t}$ can be recursively generated as follows (see also Lee (1997)).

1. Draw μ_i from a standard normal distribution.

2. For the first period,

(a) Calculate $d_{i,1} = \Psi_{i,1} + \sigma\mu_i$ (assuming the following initial conditions $\varepsilon_{i,0} = 0$ and $y_{i,0} = 0$ for all individuals)

(b) Calculate $a_{i,1} = \Phi(d_{i,1}) * I(y_{i,1} = 1) + \Phi(-d_{i,1}) * I(y_{i,1} = 0)$

(c) Calculate $b^0_{i,1} = u_{i,1} * \Phi(-d_{i,1})$

(d) Calculate $b^1_{i,1} = \Phi(-d_{i,1}) + u_{i,1} * \Phi(d_{i,1})$

(e) Calculate $e_{i,1} = \Phi^{-1}\left(b^0_{i,1}\right) * I(y_{i,1} = 0) + \Phi^{-1}\left(b^1_{i,1}\right) * I(y_{i,1} = 1)$

(f) Obtain $\varepsilon_{i,1} = e_{i,1}$

3. For $t > 1$,

(a) Calculate $d_{i,t} = \Psi_{i,t} + \gamma \sum_{j=1}^{t} \delta^{j-1} y_{i,t-j} + \sigma\mu_i + \rho\varepsilon_{i,t-1} + \nu_{i,t}$, where $\nu_{i,t}$ is drawn from a standard normal distribution

(b) Calculate $a_{i,t} = \Phi(d_{i,t}) * I(y_{i,t} = 1) + \Phi(-d_{i,t}) * I(y_{i,t} = 0)$

(c) Calculate $b_{i,t}^0 = u_{i,t} * \Phi(-d_{i,t})$

(d) Calculate $b_{i,t}^1 = \Phi(-d_{i,t}) + u_{i,t} * \Phi(d_{i,t})$

(e) Calculate $e_{i,t} = \Phi^{-1}\left(b_{i,t}^0\right) * I(y_{i,t} = 1) + \Phi^{-1}\left(b_{i,t}^1\right) * I(y_{i,t} = 0)$

(f) Obtain $\varepsilon_{i,t} = e_{i,t} + \rho\varepsilon_{i,t-1}$

This is done m times. The simulated likelihood is then

$\mathcal{L} = \sum_{i=1}^{n} ln\left\{\frac{1}{m}\sum_{j=1}^{m}\prod_{t=1}^{T_i} a_{i,t}\right\}$

Asymptotic properties of this estimator are discussed in Lee (1997) as well as in the references in that paper.

D.1.2 Ordered Outcomes

The simulated likelihood function for the dynamic ordered probit proceeds in a similar fashion but modified to accommodate the ternary nature of our outcomes. Specifically, for each independent simulation run (j), $e_{i,t}$ can be recursively generated as follows:

1. Draw μ_i from a standard normal distribution.

2. For the first period,

(a) Calculate $d_{i,1} = \Psi_{i,1} + \sigma\mu_i$ (assuming the following initial conditions $\varepsilon_{i,0} = 0$ and $c_{i,0} = 0$ for all individuals)

(b) Calculate $a_{i,1} = \Phi(\theta_1 - d_{i,1}) * I(c_{i,1} = 0) + [\Phi(\theta_2 - d_{i,1}) - \Phi(\theta_1 - d_{i,1})] * I(c_{i,1} = 1) + [1 - \Phi(\theta_2 - d_{i,1})] * I(c_{i,1} = 2)$

(c) Calculate $b_{i,1}^0 = u_{i,1} * \Phi(\theta_1 - d_{i,1})$

(d) Calculate $b_{i,1}^1 = \Phi(\theta_1 - d_{i,1}) + u_{i,1} * [\Phi(\theta_2 - d_{i,1}) - \Phi(\theta_1 - d_{i,1})]$

(e) Calculate $b_{i,1}^2 = \Phi(\theta_2 - d_{i,1}) + u_{i,1} * [1 - \Phi(\theta_2 - d_{i,1})]$

(f) Calculate $e_{i,1} = \Phi^{-1}\left(b_{i,1}^0\right) * I(c_{i,1} = 0) + \Phi^{-1}\left(b_{i,1}^1\right) * I(c_{i,1} = 1) + \Phi^{-1}\left(b_{i,1}^2\right) * I(c_{i,1} = 2)$

(g) Obtain $\varepsilon_{i,1} = e_{i,1}$

3. For $t > 1$,

(a) Calculate $d_{i,t} = \Psi_{i,t} + \gamma_1 \sum_{j=1}^{t} \delta^{j-1} 1\left(c_{i,t-1} = 1\right) + \gamma_2 \sum_{j=1}^{t} \delta^{j-1} 1\left(c_{i,t-1} = 2\right) + \sigma\mu_i + \rho\varepsilon_{i,t-1} + \nu_{i,t}$, where $\nu_{i,t}$ is drawn from a standard normal distribution

(b) Calculate $a_{i,t} = \Phi\left(\theta_1 - d_{i,t}\right) * I\left(c_{i,t} = 0\right) + \left[\Phi\left(\theta_2 - d_{i,t}\right) - \Phi\left(\theta_1 - d_{i,t}\right)\right] * I\left(c_{i,t} = 1\right) + \left[1 - \Phi\left(\theta_2 - d_{i,t}\right)\right] * I\left(c_{i,t} = 2\right)$

(c) Calculate $b^0_{i,t} = u_{i,t} * \Phi\left(\theta_1 - d_{i,t}\right)$

(d) Calculate $b^1_{i,t} = \Phi\left(\theta_1 - d_{i,t}\right) + u_{i,t} * \left[\Phi\left(\theta_2 - d_{i,t}\right) - \Phi\left(\theta_1 - d_{i,t}\right)\right]$

(e) Calculate $b^2_{i,t} = \Phi\left(\theta_2 - d_{i,t}\right) + u_{i,t} * \left[1 - \Phi\left(\theta_2 - d_{i,t}\right)\right]$

(f) Calculate $e_{i,t} = \Phi^{-1}\left(b^0_{i,t}\right) * I\left(c_{i,t} = 0\right) + \Phi^{-1}\left(b^1_{i,t}\right) * I\left(c_{i,t} = 1\right) + \Phi^{-1}\left(b^2_{i,t}\right) * I\left(c_{i,t} = 2\right)$

(g) Obtain $\varepsilon_{i,t} = e_{i,t} + \rho\varepsilon_{i,t-1}$

Similar to the binary case, this is done m times and the simulated likelihood is

$\mathcal{L} = \sum_{i=1}^{n} ln \left\{ \frac{1}{m} \sum_{j=1}^{m} \prod_{t=1}^{T_i} a_{i,t} \right\}$

Asymptotic properties of this estimator are discussed in Lee (1997) as well as in the references in that paper.

D.2 Estimates

Table D.1: Estimates from binary probits

	Model 1	Model 2	Model 3
Black	-0.079	-0.153	-0.054
	(0.064)	(0.103)	(0.074)
Hispanic	0.004	0.046	0.029
	(0.066)	(0.071)	(0.079)
Family income	0.001	0.004	0.002
	(0.003)	(0.006)	(0.005)
Mother High School	0.049	0.026	0.033
	(0.049)	(0.092)	(0.076)
Mother College	0.015	-0.015	-0.005
	(0.052)	(0.092)	(0.099)
Mother's age	-0.001	0.0003	-0.001
	(0.003)	(0.007)	(0.005)
Urban	0.037	0.095	0.053
	(0.037)	(0.029)	(0.049)
Siblings	-0.049	-0.078	-0.052
	(0.021)	(0.031)	(0.029)
$(t - t_0)$	0.077	0.180	0.088
	(0.017)	(0.019)	(0.029)
$(t - t_0)^2$	-0.005	-0.011	-0.008
	(0.001)	(0.001)	(0.002)
Constant	-1.508	-1.970	-1.571
	(0.159)	(0.235)	(0.198)

- Standard errors in parentheses.
- The remaining parameters and model descriptions are available in Table 3.4 together with likelihood values and AIC.

Table D.2: Estimates from an ordered probit polya model

	Estimate	Standard error
Black	-0.056	0.082
Hispanic	0.047	0.075
Family income	0.001	0.005
Mother High School	0.031	0.06
Mother College	-0.006	0.056
Mother's age at birth	-0.0002	0.005
Urban	0.082	0.054
Siblings	-0.05	0.026
$(t - t_0)$	0.124	0.023
$(t - t_0)^2$	-0.01	0.001

- The remaining parameters and model descriptions are available in Table 3.10 together with likelihood values and AIC.