NBER WORKING PAPER SERIES

UNEMPLOYMENT, ALCOHOL AND TOBACCO USE: SEPARATING STATE DEPENDENCE FROM UNOBSERVED HETEROGENEITY

Monica Deza

Working Paper 31453
http://www.nber.org/papers/w31453

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2023

This study is based upon work supported by the National Science Foundation under award number SES #2018438 (PI: Monica Deza). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Unemployment, Alcohol and Tobacco Use: Separating State Dependence from Unobserved Heterogeneity
Monica Deza
NBER Working Paper No. 31453
July 2023
JEL No. I12,J01

ABSTRACT

Previous literature presents mixed evidence on the effect of alcohol consumption on labor market outcomes. On one hand, heavy alcohol consumption has been shown to have detrimental effects on labor market outcomes. On the other hand, moderate consumption is positively associated with wages and employment. Despite substantial reduced form evidence, previous literature has not been able to separately identify the causal pathways linking moderate versus heavy alcohol use to labor market performance due to the lack of natural experiments that only target moderate versus heavy drinking, as well as limitations of available structural methods that model state dependence and unobserved heterogeneity. This study develops a multiple-equation dynamic discrete choice ordered logit model, which allows separate identification of the contribution of state dependence (within and between outcomes) and unobserved heterogeneity. I apply this newly-developed model to differentiate the effects of moderate and heavy drinking, after accounting for other correlated unobserved heterogeneity. This study finds that moderate alcohol use increases employment, which is consistent with moderate alcohol consumption being a venue for social capital accumulation. Policies that target alcohol consumption separately by dosage level may be beneficial to employment in ways that have not previously been expected.

Monica Deza
Department of Economics
Syracuse University
Maxwell School of Citizenship
and Public Affairs
426 Eggers Hall
Syracuse, NY 13244
and NBER
monicadeza@gmail.com

I.INTRODUCTION

Heavy drinking is common in the United States, as one in six U.S. adults binge drink and 25% binge drink at least weekly, defined as consuming 5 or more drinks on an occasion for men or 4 or more drinks on an occasion for women[2]. While the negative effect of heavy alcohol consumption is mostly unambiguous and well understood (Kenkel, 1993) particularly for socially costly outcomes such as drunk-driving and alcohol-related mortality, the effect of alcohol use on labor market outcomes- such as employment - is still ambiguous despite the vast public interest and research on the topic. Given the role of public policy over the distribution and use of alcohol, there is significant public interest to understand the external effects and social cost of alcohol use.

There are several mechanisms through which alcohol consumption has a detrimental effect on employment and earnings, particularly among heavy alcohol users (Terza, 2002; Mullahy and Sindelar, 1996; Bockerman et al, 2017; Mullahy and Sindelar,1996; Kenkel and Ribar, 1994; Fench and Zarkin, 1995; Hamilton and Hamilton, 1997). For instance, alcohol misuse is associated with being laid off, being fired, having conflicts with coworkers, having longer unemployment spells (French et al, 2011), exhibiting job withdrawal behaviors such as spending more time on nonwork activities, taking longer lunch breaks, leaving work earlier (Lehman and Simpson, 1992), higher work absenteeism (Nordstrom, 2006; Nordstrom and Moan, 2009; Johansson, Bockerman and Uutela, 2008; Laaksonen et al, 2009; Salonsalmi et al, 2009; Roche et al, 2008; Anderson, 2012) and several health problems (Mullahy and Sindelar, 1993). Contrarily, moderate alcohol consumption contributes to building social capital, which in turn has positive effects on employment and wages (Peters and Stringham, 2006; Bockerman et al, 2017). For instance, drinkers earn more than abstainers, with an additional wage premium for

[2] https://www.cdc.gov/alcohol/fact-sheets/binge-drinking.htm

social (Peters and Stringham, 2006) and moderate drinkers (French and Zarkin, 1995). In particular,-prime age workers (30-50 years old) who consume moderate levels of alcohol (as measured as individuals drinking between 2 and 3 drinks per day on average) have the highest wage premium (French and Zarkin, 1995; Heien, 1996). [3]

These patterns could be causal (e.g. if heavy alcohol consumption is detrimental to health while moderate consumption increases social capital) or they may simply be correlational relationships (e.g. if individuals with more self-control are more likely to consume alcohol in moderation and also more likely to be employed at any given point in time regardless of their alcohol consumption). Understanding the relationship between alcohol use and individual's success in the labor market and separately identifying the causal effect of moderate versus heavy alcohol consumption is crucial for policy recommendation given the high prevalence of alcohol consumption among individuals of working age and the importance of employment for economic prosperity.

How can research separately identify the effects of moderate drinking from the effects of heavy drinking? If it were possible to run a controlled experiment, one would want to randomly assign alcohol abstention, moderate alcohol consumption, and heavy alcohol consumption to three different groups of individuals, comparing subsequent employment outcomes among the three groups. As such an experiment is impractical and infeasible for a variety of reasons, it is necessary to look for another approach to answering this question. Ordinarily, a researcher might

[3] Moderate drinkers have better measures of health- as measured by lower rates of cardiovascular diseases (Marmot and Brunner, 1991; Coate, 1993) and producing higher density of lipoprotein cholesterol (Linn et al, 1993; Coate, 1993)- relative to both abstainers and heavy drinkers. This nonlinear relationship translates into a nonlinear relationship between alcohol and earnings as moderate drinkers earn more than abstainers (Barrett, 2002; French and Zarkin, 1995; Hamilton and Hamilton, 1997; Heien, 1996; Bockerman et al, 2017) or heavy users (Lee, 2003; Zarkin at al, 1998; Bockerman et al, 2017).

seek to identify a natural experiment that exogenously shifts alcohol consumption separately at the intensive as well as the extensive margin of consumption. However, while policymakers have control over a wide array of tools that influence alcohol use – for example, alcohol excise taxes, minimum legal drinking age and mandated closing times for bars among others– variation in drinking brought about by these policies does not allow us to discern between the effects of moderate versus heavy drinking. In other words, while policymakers have used tools at their disposal to elicit changes in the volume of alcohol consumed, none of these tools explicitly cause heavy drinkers to consume alcohol moderately without inducing some heavy users to abstain from alcohol. Therefore, studies that evaluate the effect of alcohol consumption induced by such policies would be estimating the aggregate effect of both moderate consumption and alcohol abstention[4].

This limitation means that natural experiments cannot be used to causally disentangle the effect of moderate versus heavy alcohol consumption on employment. In the absence of a randomized controlled trial or a natural experiment, how can researchers separately identify the effects of moderate versus heavy drinking? An alternative approach is to use structural methods, allowing an econometrician to write down a behavioral model that separates tastes from habit formation. That is, models that separate state dependence from unobserved heterogeneity (Heckman, 1981b). Similar models have been previously used to explain persistence sexual behavior among teenagers (Arcidiacono, Khwaja and Ouyang, 2011), female labor supply (Hyslop, 1999), welfare participation (Card and Hyslop, 2005) and also to separate preferences from habit formation in brand purchases in order to explain brand loyalty (Keane, 1997). Deza

[4] As an example, exogenous increases in alcohol taxes decreased ethanol consumption among heavy drinkers but this reduction is not statistically different from reductions in ethanol consumption among moderate drinkers (Gebritz et al, 2021; Saffer et al 2022).

(2015) expanded those statistical models that separated state dependence from unobserved heterogeneity for one product (e.g. models that established whether past consumption of a particular product increases the probability of consumption of that particular product) to a multi-product setting with non-mutually exclusive outcomes, in order to explain persistence as well as transitions between alcohol, marijuana and hard drugs. However, none of those models allowed the outcomes to vary by dosage level.

This study is the first one to address this gap in the literature, as it develops a structural model that separately identifies habit formation from unobserved preferences and is the first study to separately estimate the effect of moderate versus heavy alcohol consumption on an individual's probability of maintaining full-time employment. In particular, the developed multiple-equation model is flexible enough to account for reverse causality by dosage level[5], for correlated time-invariant preferences between alcohol and employment,[6] and accounts for the initial condition problem caused by the fact that their entire lifetime trajectory of alcohol

---

[5]While the main goal of the paper is to study the effect of alcohol on employment, the model takes into account reverse causality. That is, it takes into account that there are several mechanisms through which employment may increase alcohol consumption. For instance, stress in the workplace can be a risk factor for alcohol misuse given the strong association between stress and problem drinking (Anderson, 2012). This pattern is consistent with lower rates of problematic drinking during economic downturns, which are attributed mostly to income reductions (Ruhm, 1995). On the other hand, there are also several mechanisms through which employment may decrease alcohol consumption. Ruhm (2000) finds that unemployment increases alcohol consumption, particularly recreational dinking. These findings may be reconciled if economic downturns may increase recreational drinking and decrease problematic drinking (Ruhm, 1995; Ruhm, 2000), and they are consistent with documented lower measures of mental health that arise as a result of worsening labor market conditions (Charles and DeCicca, 2008). I evaluate the extent to which moderate versus heavy alcohol consumption (e.g. dosage levels for alcohol are abstention, moderate, and heavy) and part-time versus full-time employment affect each other (e.g. dosage levels for employment are not-employed, part-time and full-time).

[6] As described later in the model section, I use mass-point mixing models to account for time-invariant multidimensional unobserved heterogeneity (e.g., Heckman and Singer, 1984).

consumption and employment is not observed by the econometrician (e.g., Heckman 1981b). Finally, I control for beer taxes and local unemployment rate as the variables that enters only the latent utility for alcohol and employment, respectively, while leaving the other utility function unaffected.

Using this newly-developed multivariate *ordered* logit model and the National Longitudinal Survey of Youth 1997 (NLSY97), this paper examines employment and alcohol consumption trajectories for respondents between the ages of 18 and 22 with less than a college between years 2002 and 2011[7]. This study finds that moderate alcohol consumption increases employment, and these results remain robust to specifications that allow correlated transitory shocks between alcohol and employment to be correlated (bivariate probit model) as well as specifications that allow second lagged outcomes to affect current outcomes (bivariate ordered logit model with second-order state dependence). On the other hand, the effect of heavy alcohol consumption has a negative effect on employment, which dissipates with alternative specifications.

Relative to the existing literature, I make three main contributions. First, I extend the consideration of state dependence and unobserved heterogeneity to an *ordered* multiproduct setting, where the outcomes have dosage levels, and outcomes are not necessarily mutually exclusive. That is, I extend existing techniques for estimating trivariate logit models (Deza, 2015) to now incorporate different dosage levels. Second, this paper uses this newly-developed methodological contribution to re-evaluate the relationship between alcohol and employment

[7] The NLSY97 follows respondents starting in 1997, who were 12-16 years old at the first wave. This study places the two following restrictions on the subsample. First, I focus on years 2002-2011 since respondents are at least 18 years old in 2002, and therefore their employment trajectories starting in year 2002 are not affected by high school attendance. Second, I focus on respondents with less than a college degree at the final wave of the survey in order to prevent their employment trajectory from being affected by schooling decisions.

addressing reverse causality and correlated unobserved heterogeneity. Unlike papers that exploit natural experiments that affect overall alcohol consumption, the bivariate ordered structural model enables me to separately identify the causal effect of moderate and heavy alcohol consumption on employment, and hence complements the reduced form literature. Finally, one of the fundamental problems of economics is the ability to recommend policies. Among the main attractions of structural models is that they enable researchers to simulate policy counterfactuals, while holding preferences and other parameters constant. The third contribution of this paper is that I use the estimated parameters simulate a variety of policy counterfactuals. In particular, I evaluate the extent to which a one-time shock in moderate versus heavy alcohol consumption affects future employment. In the absence of actual policies that target alcohol separately at the intensive and extensive margin, these policy counterfactuals shed light on potential negative externalities of currently existing alcohol regulation policies.

This remainder of the paper is organized as follows. The next section discusses the data, while Section 3 discusses the model. Section 4 discusses empirical results, presents specifications diagnostics, and discusses alternative specifications as a robustness check. Section 4 also presents counterfactual experiments that simulate the extent to which simulated policies that target a particular dosage alcohol level affect employment, relative to policies that affect alcohol consumption without targeting any dosage level. Finally, Section 5 summarizes and concludes.

II.DATA

The National Longitudinal Survey of Youth 1997 (NLSY97) is a longitudinal dataset for a nationally representative sample of 8,984 respondents between the ages of 12 and 16 as of December 31st, 1996. Given that the goal of the study is to model the reinforcing patterns of employment and alcohol use over the respondents' life course, the NLSY97 is a nearly ideal dataset due to its longitudinal nature, as it provides yearly data on the respondent's alcohol consumption and employment status.

I use the detailed yearly data on alcohol consumption and its frequency of use to define three levels of alcohol consumption (alcohol users): none (non-drinker), low (moderate drinker), and high (heavy drinker). The NLSY97 asked respondents whether they consumed alcohol in the past month, the number of days of alcohol consumption, and the number of days of alcohol binge, as defined as five or more drinks. I define non-drinker as having zero drinking days in the past month, heavy drinker as using alcohol at least 20 days with at least 90 percent of those drinking days being a binge, and moderate drinker as neither non-drinker nor heavy drinker[8]. I explore with three alternative definitions as robustness checks where I define heavy drinker as using alcohol at least 20 days with at least 70 percent of those drinking days being a binge, as using alcohol at least 15 days with at least 90 percent of those drinking days being a binge, and using alcohol at least 15 days with at least 70 percent of those drinking days being a binge.

Similarly, I use the data on employment status and hours of work to define three levels of employment (worker): none (not-working), low (part-time worker), and high (full-time worker). The NLSY97 employer roster matches every respondent at every wave with every employer for whom the respondent worked since the last interview, distinguishing between whether this is an

[8] I define non-drinker as having zero drinking days in the past month, independent of whether they consumed alcohol in the past year, since restricting non-users to individuals who abstained from alcohol consumption in the past month and past year would not separately identify true lifetime abstainers from previous heavy drinkers who abstain from alcohol in response to health consequences of past heavy drinking. Previous literature indicates that long-term abstainers and current abstainers who are former drinkers face different labor market conditions (Bockerman et al, 2017; French and Zarkin, 1995; Heien, 1996).

employee-type job, freelance job, self-employment or military service[9]. Using information of whether the respondent is matched to at least one employer and the number of work hours corresponding to each employee, I determine whether the respondent worked since the last interview and compute the total number of hours worked across all reported jobs for those who reported more than one employer, which I ultimately use to determine total number of work hours. I define not-working as either not matched to an employer from the employer roster or matched to an employer but reporting zero hours of work in the past week. Part-time workers are defined as those who are matched to at least one employer from the employer roster and reporting a positive number of work hours in the past week that are at most 30. Finally, full-time workers are those that are matched to at least one employer and report working more than 30 hours in the past week.

I use the restricted version of the NLSY97, which provides state of residence at every wave, and match each respondent-year with the respective state-year beer tax and local unemployment rate, as outcome-specific controls corresponding to the latent utility of alcohol and employment, respectively[10].

From the entire sample of 8,984 adolescents interviewed in 1997, I place the following restrictions for the main subsample of interest. First, I restrict the analysis to the 5,617 respondents who were not lost due to attrition between 2002 and 2011. I focus on waves starting in 2002 since respondents would be at least of age 17 in 2002 for the initial conditions and at least of age 18 for the non-initial conditions in 2003 and after. Restricting the sample to periods that do not coincide

[9] https://www.nlsinfo.org/content/cohorts/nlsy97/topical-guide/employment

[10] Every respondent has at least one non-missing state of residence, which corresponds to the first wave. Therefore, I replace the missing state of residence with the most recently reported state of residence for that respondent. I match each respondent-year with the respective state excise tax on a gallon of beer from "History of Beverage Alcohol Tax Changes" the Distilled Spirits Council of the United States (DISCUS) and with the state-year unemployment rate from the Bureau of Labor Statistics.

with high school years prevents me from including respondents who are not in the labor force because they are still completing their high school degree. Second, from that subsample, I further restrict the analysis to the 4,648 respondents with valid information about whether they consumed alcohol, the number of days of consumption in the month prior to the interview, and the number of days of binge drinking in the month prior to the interview in all waves between 2002 and 2011. Third, among those, 4,203 respondents had a valid employment status and valid hours of work. Among respondents that are considered to be employed, they are considered to report valid hours of work if they reported a non-missing number of work hours for each of the matched employers[11]. Fourth, I restrict to the 4200 observations with non-missing demographics, which will be used as controls (indicators for whether the respondent is male, for whether there is a resident father in the household, and for the time-varying age category of the respondent). Fifth, I restrict to the 4059 respondents with information on cigarette consumption in the past month prior to interviews between 2002-2011 for the robustness section. Finally, after restricting to those with valid information on highest educational degree attained, the subsample with valid data for all relevant variables is composed of 4050 respondents. Among those, 2,781 respondents had low levels of education, as defined by their highest degree completed being at most high school, GED or and Associates Degree. The rationale behind restricting the study with individuals with less than a college degree is because I want to model the employment dynamics among individuals who are most likely to be in the labor force, and hence exclude full-time students who are seasonal workers.

[11] Given that some respondents are matched to more than one employer, I compute the total number of work hours adding up the number of work hours across matched employers. This total of number of work hours is used to determine whether any given working respondent works full-time or part-time.

There are some natural concerns about using longitudinal data. First, the study relies on respondents who were not lost due to attrition between waves 2002 and 2011, which could lead to selection issues. Second, even conditional on not being lost due to attrition, some respondents may be differentially likely to report valid answers related to their alcohol consumption or employment, which are crucial for this study[12].

Third, the self-reported nature of the NLSY raises questions about measurement error; however, the NLSY97 takes particular precaution to reduce underreporting when collecting answers to substance consumption questions by using computer-assisted self-interviews (ACASI), which reduces underreporting of risky behaviors compared to other interview methods (Brener et al, 2003). This is reflected in the fact that the NLSY97 reported rates of substance use are consistent with two non-longitudinal major datasets on drug use (i.e., the National Study of Drug Use and Health, NSDUH, and Monitoring the Future, MTF).[13]

The first column of Table 1 presents summary statistics for the overall sample of 8.984 respondents, the second column focuses on the subsample composed of those who were not lost due to attrition between 2002 and 2011 and who also reported valid values for all relevant variables between 2002 and 2011, and the third column further restricts the subsample of column 2 to respondents who are poorly educated, which are individuals with education less than a BA by the last wave. The difference across subsamples can be summarized as follows. The subsample with non-missing data (column 2) has a lower share of males, a slightly higher probability of moderate

[12] Among NLSY97 respondents not lost due to attrition, 94.6%, 92.8% and 94.5% reported valid answers to their alcohol, marijuana, and hard drugs-related questions for waves 1997-2008 (Deza, 2015). That is, missing variables related to substance use are mostly caused by respondents being lost due to attrition and not respondents purposely avoiding drug-related questions.

[13]See Deza (2015) Online Appendix B for a detailed discussion regarding comparisons between the NLSY97, NSDUH, and MTF. Table A3 of Deza (2015) compares the rates of past year, past month and lifetime substance use between NLSY97, NSDUH, and MTF.

levels of alcohol consumption in 2002, higher probability of employment particularly part-time employment, relative to the full sample. After further restricting the subsample to poorly educated respondents (column 3), there is a lower probability of having a resident father, higher probability of alcohol abstention, and higher probability of full-time employment.

Table 2 presents information on the dynamic patterns of alcohol use and employment, which illustrate the key stylized fact that motivates this study: The probability of being employed full-time is highest among those who consumed moderate levels of alcohol (63%), slightly lower by heavy-drinkers (61.9%) and significantly lowest among non-drinkers (50.1%). It is worth to emphasize that we are not defining heavy-drinkers as necessarily individuals with alcohol use disorders, but just as those who consume alcohol on a regular basis (at least 20 days in the past month) and who mostly binge on alcohol when they drink (90 percent of drinking days in the past month, conditional on being at least 20, were binge drinking). This pattern remains consistent across alternative definitions of heavy drinking.

The stylized facts in Table 2 can be summarized as follows: (i) both alcohol and employment levels exhibit serial persistence by dosage level (e.g. full-time employment is highest among those who were employed full-time in the previous period, and heavy alcohol consumption is highest among those who used heavy alcohol levels in the previous period, etc.), (ii)There is some evidence of a positive association between lag alcohol and employment for moderate levels of alcohol and negative for heavy levels of alcohol.

III. MODEL, IDENTIFICATION AND ESTIMATION

This paper develops an ordered bivariate model with the final goal of separately identifying true state dependence within and between alcohol and employment, taking into account different dosage levels. Ordered models allow me to distinguish between different dosage levels at each point in time; that is, it allows for moderate alcohol consumption to affect employment differentially from heavy alcohol consumption. Similarly, it allows for part-time employment to affect alcohol consumption differentially from heavy alcohol consumption.

I estimate a bivariate ordered logit model with three potential outcomes for both alcohol and employment in order of intensity by $\{0,1,2\}$. For alcohol use, these correspond to abstention, moderate and heavy alcohol consumption, as defined in the data section. For employment, these correspond to no employment, part-time employment and full-time employment, as defined in the data section.

The model considers an individual who maximizes his or her utility by choosing whether and how much to consume alcohol and how much to work. Let $U_{i,t}^{j}$ denote the utility that individual i experiences from engaging in behavior j[14] in year t, which are not mutually-exclusive, and assume that the utility of outcome j in period *t* depends on a outcome-specific time trend[15] $\delta_{1j}(t - t_0)$, on a set of observed characteristics of the individual X_{it}, [16] on a set of indicators by dosage level ($Y_{i,k,t-1}^{Low}, Y_{i,k,t-1}^{High}$) that equals 1 if person i engaged in low or high dosage levels of outcome k in period t-1 (j=k={alcohol,employment}), on a set of outcome-specific time-varying variables ($Z_{i,t}^{j}$)[17] that only enter the latent utility for outcome j, while being excluded from

[14] Let j={a,e} for alcohol and employment, respectively

[15] The variable t is the year that corresponds to the utility $U_{i,t}^{j}$, while t_0 is always 2002.

[16] The vector X_{it} represents time-invariant characteristics (gender and whether the individual comes from a single-headed household) and time-varying observable characteristics (age) of consumer i in time t.

[17] These outcome-specific time-varying variables are state-specific beer taxes for alcohol and the state-specific unemployment rate for employment.

entering the latent utility for other outcomes j' when $j \neq j'$,and on a combination of a permanent unobserved taste component α_i^j and a transitory unobserved component $\varepsilon_{i,t}^j$.[18]

The latent utilities for each outcome differentiate low lagged dosage levels from high lagged dosage levels[19] and are specified as follows:

$$U_{ijt} = \underbrace{\delta_{1j}(t-t_0) + X_{it}\beta_j + \sum_{k=1}^{J}\gamma_{kj}^{Low}Y_{i,k,t-1}^{Low} + \sum_{k=1}^{J}\gamma_{kj}^{High}Y_{i,k,t-1}^{High} + \varphi_j Z_{ijt} + \alpha_{ij}}_{V_{ijt(\alpha_{ij})}} + \varepsilon_{ijt} \qquad (1)$$

At each period, Y_{ijt} has a value of 0, 1 or 2 if the following equations hold, respectively.

$$Y_{ijt} = \begin{cases} 0, & U_{ijt} \leq c_{1j} \\ 1, & c_{1j} \leq U_{ijt} \leq c_{2j} \\ 2, & U_{ijt} \geq c_{2j} \end{cases} \qquad (2)$$

This model estimates the thresholds c_{1j}, and c_{2j}, along with the parameters of the latent utilities $V_{ijt}(\alpha_j)$. With a logistic error assumption, the probability of an individual choosing outcome j with dosage level 0,1 or 2 has a closed form solution and can be written as follow

$$P\left(Y_{ijt} = 0 \middle| \overline{Y}_{i,t-1}, \alpha_{ij}\right) = \frac{\exp\left[c_1 - V_{ijt(\alpha_{ij})}\right]}{1+\exp\left[c_1 - V_{ijt(\alpha_{ij})}\right]} \qquad (3)$$

$$P\left(Y_{ijt} = 1 \middle| \overline{Y}_{i,t-1}, \alpha_{ij}\right) = \frac{\exp\left[c_2 - V_{ijt(\alpha_{ij})}\right]}{1+\exp\left[c_2 - V_{ijt(\alpha_{ij})}\right]} - \frac{\exp\left[c_1 - V_{ijt(\alpha_{ij})}\right]}{1+\exp\left[c_1 - V_{ijt(\alpha_{ij})}\right]} \qquad (4)$$

$$P\left(Y_{ijt} = 2 \middle| \overline{Y}_{i,t-1}, \alpha_{ij}\right) = 1 - \frac{\exp\left[c_2 - V_{ijt(\alpha_{ij})}\right]}{1+\exp\left[c_2 - V_{ijt(\alpha_{ij})}\right]} \qquad (5)$$

Since we do not observe employment and alcohol use for the entire lifetime of the respondent, I am unable to construct the likelihood function for all years since the stochastic process of employment and alcohol choices started. Therefore, the next step is to separately the

[18] The unobserved heterogeneity specification allows for the permanent preferences for each drug to be arbitrarily correlated across outcomes.
[19] The latent utility for the ordered logit model does not include an intercept because the intercept would not be identified separately from the threshold. Also, the random effect affects the intercept, and allows the threshold to vary by type.

initial conditions equations for each outcome j (e.g., Keane, 1997; Heckman 1981b). Given that the latent utility for the initial period does not have lagged values, I need a separate specification for the outcomes of the "initial conditions" for year 2002, where I allow for the coefficients on X_{i0} and Z_{ij0} , and the threshold parameters $\{c_{1,0j}\ c_{2,0,j}\}$ in the initial conditions to differ from those in later periods $t > t_0$.[20]

$$U_{ij0} = \underbrace{X_{it}\lambda_{0j} + \varphi_{j,0} Z^j_{i,t} + \alpha^j_{i,0}}_{V^j_{i,t}(\alpha^j_{i,0})} + \varepsilon^j_{i,t} \tag{6}$$

Next, I specify the distribution of the unobserved heterogeneity to allow for correlation between the unobservable time-invariant preferences for outcomes j and also between $t = t_0$ and $t > t_0$. In order to account for such correlation, I use mass-point mixing models to account for time-invariant multidimensional unobserved heterogeneity (e.g., Heckman and Singer, 1984). That is, I specify that α^j_i and $\alpha^j_{i,0}$ come from a mass point distribution that describes the time-invariant unobserved heterogeneity for 3 different types of individuals. This approach treats $\alpha^a_i, \alpha^e_i, \alpha^h_i$, $\alpha^a_{i,0}, \alpha^e_{i,0}$ as random variables that are highly correlated in such a way that when α^j_i comes from type p, then $\alpha^j_{i,0}, \alpha^{j'}_{i,0}$ and $\alpha^{j'}_i$ (for $j \neq j'$) also come from type p. That is, every individual i of type "p" shares a vector of unobserved heterogeneity $\alpha_p = (\alpha^a_p, \alpha^e_p, \alpha^a_{p,0}, \alpha^e_{p,0})$ for type p, p={1,2,3}.[21] The

[20] Failure to separately model the initial conditions leads to biased and inconsistent coefficients in the presence of serial correlation (e.g., Heckman, 1981b) and when the stochastic process of the relevant outcome started prior to observed periods (Altonji et al, 2010). On the other hand, studies where the respondents have identical outcomes in the periods prior to the observed sample do not require to separately model the initial conditions problem (e.g., Card and Hyslop 2005).

[21] I normalized the random effect for Type 1 to zero in order to separately identify the thresholds. That is, the random effects allow the outcome-specific intercepts c_{1j} and c_{2j} to vary by type. I specify the vector of unobserved heterogeneity $\alpha_p = (\alpha^a_p, \alpha^e_p, \alpha^a_{p,0}, \alpha^e_{p,0})$ for type p, p={1,2,3} as follows: Type 1 $(0,0,0,0)$, Type 2 $(\alpha^a_2, \alpha^e_2, \alpha^a_{2,0}, \alpha^e_{2,0})$, and Type 3 $(\alpha^a_3, \alpha^e_3, \alpha^a_{3,0}, \alpha^e_{3,0})$.

type is known to the individual but unknown to the econometrician[22]. This specification estimates both the location of the mass points, α_p^j and $\alpha_{p,0}^j$, and the share of individuals who belong to each type (the unconditional probability), π_p[23]. Finally, $\varepsilon_{i,t}^j$ represent transitory taste shocks, which are assumed to be drawn independently from an Extreme Value Type 1 distribution for the bivariate ordered logit model.

The main parameters of interest are γ_{kj}^{Low} and γ_{kj}^{High}, which represent the state dependence parameters. For k=j these are the effects of lagged outcome j by dosage level on current outcome j (state dependence *within* outcome). For k≠j these are the effects from past alcohol use by dosage level on employment or the effects from past employment by dosage level on alcohol use (state dependence *between* outcomes, from outcome k to outcome j).

The likelihood of a sequence of drug j consumption indicators represents the individual contribution to the entire likelihood function for an individual of type m, where the outcome Y_{ijt} has three values: {0,1,2}.

$$L_i^j(Y_{ij0,\dots,}Y_{ijT}|\alpha_p^j) = \prod_j P\left(Y_{i,2002}^j|\alpha_{p,0}^j\right) * \prod_{t=2003}^{2011} \prod_j P\left(Y_{i,t}^j|\overline{Y}_{i,t-1}, \alpha_p^j\right) \quad (7)$$

At each time period t, individual i chooses the dosage level of outcome j depending of where the utility attained from outcome j falls relative to thresholds c_{1j} and c_{2j}. This decision is made independently for alcohol and employment, conditional on the correlated unobserved heterogeneity. With the assumption that the errors are drawn from the extreme value Type 1

[22] Please refer to Heckman and Singer (1984) for a thorough discussion of the mass-point distribution

[23] I specify the unconditional probability of being type p as $\pi_p = \frac{\exp(\phi_p)}{1+\exp(\phi_2)+\exp(\phi_3)}$, where I normalize ϕ_1 to zero and the model estimates parameters ϕ_2 and ϕ_3. This restriction is solely imposed to ensure that π_p is between zero and one, and that $\sum_{p=1}^{3} \pi_p = 1$.

distribution, the probability of an individual using drug j at any given period has a closed form solution as in equation 3,4 and 5.

Finally, the individual contribution to the likelihood is the weighted average of type-specific likelihood contributions, using the unconditional probabilities, π_p, as weights.

$$L_i = \sum_{p=1}^{3} \pi_p \{\prod_j P(Y_{i,2002}^{j} | \alpha_{p,0}^{j}) * \prod_{t=2003}^{2011} \prod_j P(Y_{i,t}^{j} | \overline{Y}_{i,t-1}, \alpha_{p}^{j})\} \quad (8)$$

With 3 mass points there are 50 parameters to estimate: ten utility parameters$\{ \delta_{1j}, \beta_{1,j}, \beta_{j2,}, \beta_{3,j}, \beta_{4,j,} \beta_{5,j}, \gamma_{a,j}^{Low}, \gamma_{e,j}^{Low}, \gamma_{a,j}^{High}, \gamma_{e,j}^{High}, \}$ for each outcome j and two outcome-specific covariates $\{\varphi_1, \varphi_2\}$ for periods $t > t_0$, four utility parameters for each outcome $\{\lambda_{1,0j}, \lambda_{2,0j}, \lambda_{3,0j}, \lambda_{4,0j}\}$ and two outcome-specific covariates $\{\varphi_{1,0}, \varphi_{2,0}\}$ for the initial period ($t = t_0$), four unobserved heterogeneity parameters for each outcome $\{\alpha_2^j, \alpha_3^j, \alpha_{2,0}^j, \alpha_{3,0}^j\}$, two type-associated probability parameters $\{\phi_2, \phi_3\}$ and four threshold parameters for each outcome j $\{c_{1j}$, c_{2j}, $c_{1,0,j}$, $c_{2,0,j}\}$.

IV.EMPIRICAL RESULTS

A. Parameter Estimates

The columns of Table 3 present the parameter estimates of equation 1, which specifies the latent utilities for alcohol and employment, respectively, where the parameter γ_{kj}^{Low} (γ_{kj}^{High}) indicates the coefficient of lag low (high) levels of outcome k on the latent utility of outcome j.

The coefficients of Panel A correspond to the main subsample of interest, which is composed by the 2,781 respondents with less than a college degree[24]. The coefficients of Panel A indicate

[24] As mentioned earlier, I focus on respondents with at most high school, GED or and Associates Degree. Since college students may be more likely to enter the labor force seasonally, excluding

that both alcohol and employment exhibit state dependence *within* outcome, which increases in magnitude with dosage levels. That is, $\gamma_{jj}^{Low} > 0$, $\gamma_{jj}^{High} > 0$, and $\gamma_{jj}^{High} > \gamma_{jj}^{Low}$. In particular, $\gamma_{a,a}^{Low}$ =1.38 (SE=0.04), $\gamma_{a,a}^{High}$ =4.05 (SE=0.17), $\gamma_{e,e}^{Low}$ =0.67 (SE=0.04), $\gamma_{e,e}^{High}$ =1.97 (SE=0.04). The coefficients associated with the state dependence *between* outcomes (γ_{kj}^{Low} and γ_{kj}^{High} when $k \neq j$) show interesting patterns in support of the importance of dosage levels. The most salient finding is that lagged moderate consumption of alcohol increases future employment while high levels of alcohol decrease future employment. That is, $\gamma_{a,e}^{Low} > 0$, $\gamma_{a,e}^{High} < 0$. In particular, $\gamma_{a,e}^{Low}$ =0.14 (SE =0.04), $\gamma_{a,e}^{High}$ =-0.28 (SE =0.14). On the other hand, the coefficients do not suggest a statistically significant effect from lagged employment to future alcohol use. In particular, $\gamma_{e,a}^{Low}$ =0.03 (SE =0.06), $\gamma_{e,a}^{High}$ = 0.06 (SE =0.05).

These patterns remain consistent across different subsamples. Panel B estimates the effect of the entire sample of 4,050 respondents regardless of their terminal educational attainment. While using all respondents regardless of their educational status allows me to estimate the coefficients using a larger sample, estimating the coefficients without restricting the subsample to those with less than a college degree may not provide an accurate description of employment trajectories as some of these respondents may still be full-time students during some periods and this model does not account for education decisions on a separate equation. However, since respondents are between 18 and 23 at the initial conditions, only a subset of them are still acquiring education even without any educational restrictions. Unsurprisingly, the results remain largely unchanged as state dependence *within* outcome is positive and statistically significant for both alcohol and

them enables me to model employment dynamics among individuals who are most likely to be in the labor force consistently.

employment, and increase in magnitude with dosage levels. That is, $\gamma_{\mathrm{a,a}}^{\mathrm{Low}} = 1.473$ (SE = 0.034), $\gamma_{\mathrm{a,a}}^{\mathrm{High}} = 4.312$ (SE = 0.159), $\gamma_{\mathrm{e,e}}^{\mathrm{Low}} = 0.767$ (SE = 0.034) , $\gamma_{\mathrm{e,e}}^{\mathrm{High}} = 2.234$ (SE = 0.033). Similarly, the coefficients suggest that low (high) lagged alcohol use increases (decreases) employment ($\gamma_{\mathrm{a,e}}^{\mathrm{Low}} = 0.1799$ (SE = 0.033), $\gamma_{\mathrm{a,e}}^{\mathrm{High}} = -0.215$ (SE = 0.1275)) while lagged employment does have a statistically significant effect on alcohol consumption ($\gamma_{\mathrm{e,a}}^{\mathrm{Low}}$ =0.0032(SE =0.0566), $\gamma_{\mathrm{e,a}}^{\mathrm{High}}$ = 0.0419 (SE =0.0478)).

A potential limitation of comparing moderate alcohol users to abstainers is that some abstainers may be formerly diagnosed with alcohol misuse disorder and stopped drinking entirely at some point. Given that I define abstainers as individuals how did not drink in the month prior to the interview but may have consumed alcohol since the last interview, former alcohol misusers would only be a subset from abstainers. I re-estimate the analysis after dropping those who are likely to be recovering alcohol misusers. That is, those who consumed alcohol at least once a year in every wave up to a certain point, but reported not consuming alcohol at all during the entire year ever since the stopping point. For example, an individual who consumed alcohol every single wave until 2005, but then starting 2006 onwards reported never consuming alcohol in the entire year would be considered "likely former alcohol misuser." I re-estimate the analysis after dropping observations that went through that dramatic change of stopping alcohol consumption entirely at any point (Panel C) and those who went through that change at any point between the waves of interest 2002-2011 (Panel D). The results remain largely unchanged.

Finally, I explore with an alternative definition of heavy drinker, which is now defined as those who consumed alcohol at least 20 days in the last month but at least 70 percent of those days involve binge drinking (Panel E). In addition, I explore with a less restrictive definition where heavy drinkers are defined as those who consume alcohol at least 15 days in the last month but at least 70% or 90% of those involve binge drinking and the results remain largely unchanged (Table A5) and the results remain unchanged to alternative and less restrictive measures of heavy alcohol consumption.

The outcome-specific variables (Z_{ijt}) are required to vary at the outcome and year level, in order to be a source of identification. Even though alcohol prices would be a natural choice,[25] they do not exhibit enough variation over time. On the other hand, beer taxes exhibit enough state-year variation, they are less subject to measurement error than alcohol prices (Cook, 1981), and their negative impact on alcohol consumption has been well documented (Gebritz et al, 2021; Dave and Kaestner, 2002; Dee and Evans, 2003; Grossman et al, 1993; Lenke, 1993; Cook and Moore, 1999; Dee, 1999). Unsurprisingly, beer taxes negatively enter the latent utility of alcohol consumption (See Table A1) and this coefficient is statistically significant at the conventional level. Similarly, state-level unemployment rate negatively enters the utility employment, and this coefficient is statistically significant at the conventional level as well (Table A1).

Another important fact from the entire set of coefficients is the relationship between the location parameters for of the unobserved heterogeneity for $t > t_0$, α_p^j ,for each mass point p={1,2,3} across outcomes, which indicates that there is a share of the subsample (type 2) with a

[25] Alcohol prices (as measured by the series of price of malt beverages per 16 oz., provided by the Bureau of Labor Statistics by region and month) exhibit limited variation over time. When presented with limited over-time variation in prices, previous structural work has used non-price variables as a choice for covariates that enter the utility of one outcome while leaving other utilities unaffected. For instance, Gentzkow(2007) used respondent's internet access and usage for work and educational purposes as the exclusion restriction for the latent utility of online newspapers (Gentzkow, 2007) and Deza(2015) used state-year deviations from national trends in substance treatment center admissions for alcohol, marijuana and hard drugs as the outcome-specific covariates for the latent utility of alcohol, marijuana and hard drugs, respectively.

relatively high preference for alcohol but has a lower preference for employment ($\alpha_p^a > 0, \alpha_p^e < 0$), while another share of the subsample (type 3) with a relatively higher preference for alcohol only but same preference for employment ($\alpha_p^a > 0,\ \alpha_p^e = 0$) , relative to individuals of type 1 (See Table A1 for coefficients)[26]. These patterns are observed also in the location parameters for the unobserved heterogeneity at the initial conditions. The estimates indicate the share of individuals that are of type p (π_p) is 22%, 14%, and 64% for type 1, 2, and 3, respectively.[27]

Overall, the estimates support a story of moderate alcohol consumption increasing the probability of employment while heavy alcohol consumption decreasing the probability of employment. Similarly, past alcohol consumption, particularly in higher dosage, increases the latent utility of alcohol consumption. The same pattern holds for employment, as the probability of employment increases with employment in the past period, particularly with lag full-time employment. Finally lagged employment does not affect future alcohol consumption.

B. Quantifying the Effects of True State Dependence and Stepping-Stone Effects

The structural estimates are useful to reveal the sign of the state dependence and stepping-stone effects within and between alcohol and employment but they cannot directly be interpreted

---

[26] In particular, the vector of permanent unobserved heterogeneity for type p, $\alpha_p = (\alpha_p^a, \alpha_p^e, \alpha_{p,0}^a, \alpha_{p,0}^e)$, is the following for each type p, p={1,2,3}: Type 1 (0,0,0), Type 2 (0.91, -1.77, 1.16, -1.72), and Type 3 (2.51, 0.03, 2.38, 0.17). See Table A1 for the entire set of coefficients and see Figure 1 for the graphical display of the location parameters, which shows that Type 2 respondents have a higher (lower) permanent preference for alcohol (employment) than Type 1 respondents. On the other hand, Type 3 respondents have a higher (similar) permanent preference for alcohol (employment) relative to Type 1.

[27] The coefficients $\phi_1 = 0, \phi_2 = -0.48$ and $\phi_3 = 1.07$ presented in Table A1 enter the unconditional probability of being type p as $\pi_p = \frac{\exp(\phi_p)}{1+\exp(\phi_2)+\exp(\phi_3)}$ and result in π_1 =0.22, π_2 =0.14, and π_3 =0.64.

to answer the extent to which the probability of having a full-time job increases if the respondent consumes alcohol in moderate amounts, holding preferences constant. This section quantifies the extent to which a lagged behavior affects the probability of future behavior, by dosage level.

A main goal of this study is to compute the average partial effects (APEs) using the parameter estimates to quantify the change in predicted probability of full-time employment that is driven by lagged moderate levels of alcohol consumption relative to heavy levels of alcohol consumption, holding preferences constant: $\mathrm{P}[\mathrm{Y}_{\mathrm{i,e,t}}^{\mathrm{High}} = 1 | \mathrm{Y}_{\mathrm{i,a,t-1}}^{\mathrm{Low}} = 1]$ - $\mathrm{P}[\mathrm{Y}_{\mathrm{i,e,t}}^{\mathrm{High}} = 1 | \mathrm{Y}_{\mathrm{i,a,t-1}}^{\mathrm{High}} = 1]$. That is, the causal effect of lagged moderate alcohol consumption on current full-time employment. Given that lagged alcohol consumption for each individual-year can only be observed in one possible scenario ($\mathrm{Y}_{\mathrm{i,a,t-1}}^{\mathrm{None}} = 1$ or $\mathrm{Y}_{\mathrm{i,a,t-1}}^{\mathrm{Low}} = 1$, or $\mathrm{Y}_{\mathrm{i,a,t-1}}^{\mathrm{High}} = 1$), observed data would only allow me to compute either $\mathrm{P}[\mathrm{Y}_{\mathrm{i,e,t}}^{\mathrm{High}} = 1 | \mathrm{Y}_{\mathrm{i,a,t-1}}^{\mathrm{Low}} = 1]$ or $\mathrm{P}[\mathrm{Y}_{\mathrm{i,e,t}}^{\mathrm{High}} = 1 | \mathrm{Y}_{\mathrm{i,a,t-1}}^{\mathrm{High}} = 1]$ or $\mathrm{P}[\mathrm{Y}_{\mathrm{i,e,t}}^{\mathrm{High}} = 1 | \mathrm{Y}_{\mathrm{i,a,t-1}}^{\mathrm{None}} = 1]$.

A main advantage of dynamic discrete choice models is that I can hold all parameters constant and artificially set lagged alcohol consumption to moderate for everyone to compute $P[\mathrm{Y}_{\mathrm{i,e,t}}^{\mathrm{High}} = 1 | \mathrm{Y}_{\mathrm{i,a,t-1}}^{\mathrm{Low}} = 1]$, and then I can artificially set lagged alcohol consumption to heavy consumption for everyone to compute $\mathrm{P}[\mathrm{Y}_{\mathrm{i,e,t}}^{\mathrm{High}} = 1 | \mathrm{Y}_{\mathrm{i,a,t-1}}^{\mathrm{High}} = 1]$ .Then, I would be able to estimate this difference in probabilities separately for each respondent and period and then take the average over individuals and all periods after the initial period to compute the following APE among others.

$$\mathrm{P}[\mathrm{Y}_{\mathrm{i,e,t}}^{\mathrm{High}} = 1 | \mathrm{Y}_{\mathrm{i,a,t-1}}^{\mathrm{Low}} = 1] - \mathrm{P}[\mathrm{Y}_{\mathrm{i,e,t}}^{\mathrm{High}} = 1 | \mathrm{Y}_{\mathrm{i,a,t-1}}^{\mathrm{High}} = 1] \tag{9}$$

The APE from the trivariate ordered logit models presented in Table 4 can be summarized as follows. First, the estimates indicate that if all respondents, holding their preferences constant, were to artificially be employed full-time in any given period, their probability of full-time employment in the next period would increase by 41 percentage points relative to the counterfactual where they would be artificially not employed in any given period ($P[Y^{High}_{i,e,t} = 1|Y^{High}_{i,e,t-1} = 1]$ - $P[Y^{High}_{i,e,t} = 1|Y^{None}_{i,e,t-1} = 1]$=0.41). (Panel A Table 4). Given the literature on state dependence on employment, it is unsurprising that $P[Y^{High}_{i,e,t} = 1]$ more than doubles as I artificially change the lagged employment outcome from no-employment to full-time employment (from 0.31 to 0.73)

Second, the model predicts that the probability of heavy alcohol use is largest in response to artificially changing lagged alcohol use to heavy consumption for everyone ($P\left[Y^{High}_{i,a,t} = 1\middle|Y^{High}_{i,a,t-1} = 1\right] = 0.12$), followed by changing lagged alcohol use to moderate amounts ($P[Y^{High}_{i,a,t} = 1|Y^{Low}_{i,a,t-1} = 1] = 0.01$), and lowest by changing alcohol use to none for everyone ($P[Y^{High}_{i,a,t} = 1|Y^{None}_{i,a,t-1} = 1]$ =0.003). These results are also unsurprising given the literature on state dependence of alcohol consumption. It is worth to note that the raw data indicates that the probability of heavy alcohol consumption is 25.6 percent among those who consume heavy alcohol in the previous period, but this is heavily driven by unobserved preferences since that number drops to 12 percent once I artificially change the lag consumption to alcohol to heavy levels for everyone (Panel B Table 4)[28].

Third, Panel C of Table 4 presents the motivation for this study. The APE indicates that lagged moderate alcohol increases the probability of full-time employment by 7.8 and 2.1 percentage points relative to heavy alcohol use and alcohol abstention, respectively. That is, the probability of full-time employment increases from 50.71 percent to 58.44 percent in response to artificially

[28] While the parameter estimates simulate the patterns of full-time employment closely (Panel A and C, column 2), the model does not simulate patterns of heavy alcohol consumption as closely, and therefore their APE must be interpreted with caution.

changing lagged alcohol use from heavy to moderate for everyone. Similarly, the probability of full-time employment increases from 56.34 percent to 58.5 percent in response to artificially changing lagged alcohol use from no alcohol to moderate consumption. In a nutshell, moderate levels of alcohol consumption leads to higher rates of full-time employment than abstention or heavy use.

In a nutshell, the APE show evidence of *state dependence within outcomes*, and such state dependence increases with dosage level for both employment and alcohol (e.g. lag full-time employment increases the probability of full-time employment more than part-time employment, and lag part-time employment increases the probability of full-time employment more than lag no employment). On the other hand, the APE for *state dependence between outcomes* has nonlinear results as the probability of full-time employment increases in response to lag moderate alcohol consumption relative to both alcohol abstention and heavy alcohol use, holding all parameters constant.

C. **Robustness Checks and Specification Diagnostics**

The trivariate ordered logit model with fist-order state dependence relies on several assumptions about the transitory unobservable component (ε_{ijt}), regarding the mean E(ε_{ijt})=0, variance $E\left[\varepsilon_{ijt}^2\right] = 1$, serial correlation $E\left[\varepsilon_{ijt}, \varepsilon_{ijt-k}\right] = 0$ for k=1,2,3,4,5, contemporaneous correlation $E\left[\varepsilon_{ijt}, \varepsilon_{ij't}\right] = 0$, and cross-period correlation between outcomes, respectively $E\left[\varepsilon_{ijt}, \varepsilon_{ij't'}\right] = 0$.

In order to diagnose misspecification, I compute the sample-analogue generalized residuals ($r_{i,t}^{j}\left(\alpha_i^j\right)$) to evaluate the extent to which the assumptions about the transitory unobservable component ε_{ijt} hold. That is, I extend the sample-analogue generalized residuals previously implemented for the single binary variable setting (Card and Hyslop, 2005) and for the multiproduct binary variable setting Deza (2015) to the ordered model setting. If the model is correctly specified, those conditions must hold at the true location parameter of the unobserved heterogeneity, where $r^j_{i,t}(\alpha^j_i)$ is a generalized residual[29].

I compute the following sample-analogue generalized residuals for each individual and compare whether the sample-analogues in the following conditions are close to their expected value under the null hypothesis of a correctly specified mode in order to diagnose misspecification.

$$E[r^j_{i,t}(\alpha_i)] = 0 \tag{9}$$

$$E[r^j_{i,t}(\alpha_i)^2] = 1 \tag{10}$$

$$E[r^j_{i,t}(\alpha_i), r^j_{i,t-k}(\alpha_i)] = 0 \text{ for k=1,2,3,4,5} \tag{11}$$

$$E[r^j_{i,t}(\alpha_i), r^{j'}_{i,t}(\alpha_i)] = 0 \tag{12}$$

$$E[r^j_{i,t}(\alpha_i), r^{j'}_{i,t-k}(\alpha_i)] = 0 \text{ for k=1,2,3,4,5} \tag{13}$$

Following the literature generalized residuals for nonlinear models (Gourieroux et al, 1987; MacCall, 1994; Chesher and Irish, 1987) and the more recent literature on sample-analogue generalized residuals (Card and Hyslop, 2005; Deza, 2015), I compute the sample-analogue generalized residual ($r^j_{i,t}(\alpha^j_i)$) as the difference between the actual binary variable $Y_{ijt} = \{0,1,2\}$ and its expected value $E[Y_{ijt}|\alpha^j_i]$, normalized to correct for heteroskedasticity[30].

[29] A challenge to compute a sample-analogue of those five conditions is that the random effect is unknown to the econometrician. See Appendix B for a thorough description of the steps to constructing the generalized residual ($r^j_{i,t}(\alpha^j_i)$) at an assigned (and unknown to the econometrician) random effect α_i.

[30] Please refer to Card and Hyslop (2005) for a thorough description of sample-analogue generalized residuals and Deza (2015) for the extension to a multi-product setting with three binomial variables. This study further extends the sample-analogue generalized residuals

Given that the type is unknown to the econometrician, I compute the generalized residual in three steps, which are described in detail in Appendix B: First, I evaluate the generalized residual at each mass point, and obtain $r^{j}_{i,t}(\alpha^{j}_{1})$, $r^{j}_{i,t}(\alpha^{j}_{2})$, and $r^{j}_{i,t}(\alpha^{j}_{3})$, for type 1,2, and 3, respectively. Second, I compute the posterior probability of being type p using the observed sequence of outcomes. Finally, I compute a weighted average of the following conditions evaluated at each mass point, using the posterior probabilities as weights.

The sample-analogue generalized residuals corresponding to the bivariate ordered logit model (column 1,2 Table 5) are used a benchmark to diagnose misspecification and can be summarized as follows. First, the mean is statistically indistinguishable from zero, $E\left[r^{j}_{i,t}(\alpha_i)\right] = 0$. In particular, $E\left[r^{a}_{i,t}(\alpha_i)\right]$ =-0.001 (SE= 0.005) and $E\left[r^{e}_{i,t}(\alpha_i)\right]$ =-0.005 (SE= 0.006). Second, the variance is statistically indistinguishable from one, $E\left[r^{j}_{i,t}(\alpha_i)^2\right] = 1$. That is, $E\left[r^{a}_{i,t}(\alpha_i)^2\right]$ =0.995 (SE= 0.014) and $E\left[r^{e}_{i,t}(\alpha_i)^2\right]$ =1.049 (SE= 0.014). Third, the predicted errors have small but statistically significant serial correlation $E\left[r^{j}_{it}(\alpha_i), r^{j}_{i,t-k}(\alpha_i)\right]$ for k=1,2,3,4,5. For instance, $E\left[r^{a}_{it}(\alpha_i), r^{a}_{i,t-1}(\alpha_i)\right]$ =-0.03 *(SE*= 0.01). In order to partially relax the serial correlation assumption, I estimate a trivariate logit model with second-order state dependence[31]. Fourth, there is evidence of correlated contemporaneous predicted errors between outcomes that are, albeit small, statistically distinguishable from zero, as $E\left[r^{a}_{i,t}(\alpha_i), r^{e}_{i,t}(\alpha_i)\right] = 0.03$ (SE= 0.01). In order to address this misspecification in the bivariate ordered logit model, I estimate a bivariate

---

specification diagnostics to an ordered model where $r^{j}_{it}(\alpha_i)$ is defined as follows and Appendix B describes its construction in more detail:

$$r^{j}_{it}(\alpha_i) = \frac{Y^{j}_{i,t} - E[Y_{ijt}|\alpha^{j}_{i}]}{\sqrt{var(Y_{ijt})}}$$

[31] See Appendix C for a more detailed description of the bivariate logit model with second-order state dependence.

ordered probit model, which allows for correlated transitory shocks between employment and alcohol[32]. Finally, the cross-time, cross-outcome sample-analogue generalized residual $E\left[r_{i,t}^{j}(\alpha_i), r_{i,t-k}^{j'}(\alpha_i)\right]$ where k=1,2,3,4,5 and j,j'={alcohol, employment}support the evidence that the assumption of uncorrelated cross-time cross-outcome errors does not pose a threat to misspecification.

Taken together, I address the two potential threats to misspecification (equations 11 and 12) estimating a trivariate ordered logit model with second order state dependence and a trivariate ordered probit model with first order state dependence, respectively. When estimating the trivariate ordered logit model with second order state dependence, I estimate a variant of equation (1) where I add second order lags to the latent utility in equation[33] and model the initial conditions as described in Appendix C. When estimating the trivariate ordered probit model with first-order state dependence, I estimate a variant of equation (1) where I estimate an additional parameter that allows the transitory shocks to be correlated[34], as describe in Appendix D.

Reassuringly, the findings are consistent across models that allow for higher order state dependence and correlated transitory shocks, and can be summarized as follows. First, there is state dependence within alcohol and employment ($\gamma_{a,a}^{Low} > 0,\ \gamma_{e,e}^{Low} > 0, \gamma_{a,a}^{High} > 0, \gamma_{e,e}^{High} > 0$)

---

[32] See Appendix D for a more detailed description of the bivariate probit model with first-order state dependence.

[33] In particular, I add the following to the latent utility in equation (1)

$$\sum_{k=1}^{J} \gamma_{kj}^{2Lag_{Low}} Y_{i,k,t-1}^{Low} + \sum_{k=1}^{J} \gamma_{kj}^{2Lag_{Low}} Y_{i,k,t-1}^{High}$$

[34] In particular, the bivariate probit model with first-order state dependence estimates the additional parameter ϕ_{12}, which enters the covariance function as follows in order to guarantee the covariance is between -1 and 1

$$\mathrm{cov}(\varepsilon_{it}^{a}, \varepsilon_{it}^{e}) = -1 + 2 * \frac{\exp(\phi_{12})}{1 + \exp(\phi_{12}) + \exp(\phi_{12})}$$

which increases with dosage level ($\gamma_{a,a}^{Low} < \gamma_{a,a}^{High}$, $\gamma_{e,e}^{Low} < \gamma_{e,e}^{High}$). Second, lagged moderate consumption of alcohol increases future employment ($\gamma_{a,e}^{Low} > 0$). Third, the effect of heavy alcohol consumption on employment is ambiguous, as both the bivariate logit model with first-order state dependence (columns 1-2) as well as second-order state dependence (columns 5-6) indicate that heavy alcohol reduces employment, while the bivariate probit model with first-order state dependence (columns 3-4) indicate null effects. Fourth, employment does not have a statistically significant effect on alcohol, independent of dosage level ($\gamma_{ae}^{Low} = 0, \gamma_{ae}^{High} = 0$). Fifth, there is a positive and statistically significant effect from the second order lag ($\gamma_{aa}^{2Lag_{Low}} > 0, \gamma_{aa}^{2Lag_{High}} > 0, \gamma_{e,e}^{2Lag_{Low}} > 0, \gamma_{e,e}^{2Lag_{High}} > 0$, which increases by dosage level ($\gamma_{aa}^{2Lag_{High}} > \gamma_{a,a}^{2Lag_{Low}}, \gamma_{e,e}^{2Lag_{High}} > \gamma_{e,e}^{2Lag_{Low}}$), but they are smaller in magnitude than the first-order lagged effect ($\gamma_{aa}^{2Lag_{High}} < \gamma_{a,a}^{High}, \gamma_{ee}^{2Lag_{High}} < \gamma_{ee}^{High}, \gamma_{aa}^{2Lag_{Low}} < \gamma_{a,a}^{Low}, \gamma_{ee}^{2Lag_{Low}} < \gamma_{ee}^{Low}$). Finally, the probit model indicates that the structural residuals are positively correlated and estimates $cov(\varepsilon_{it}^{a}, \varepsilon_{it}^{e}) = 0.039$.[35] While there are benefits and limitations to the set of assumptions required by each of these three models, it is reassuring that the findings are consistent across these three specifications.

---

[35] The correlation of the generalized residuals $E\left[r_{it}^{j}(\alpha_i), r_{it}^{j'}(\alpha_i)\right]$ does not need to match the correlation of the structural residuals $cov\left(\varepsilon_{it}^{j}, \varepsilon_{it}^{j'}\right)$ in the probit model, as simulated data shows that the correlation of generalized residuals is always lower than the correlation of structural residuals (Deza, 2015). While the correlation of generalized residuals is not relevant on its own, it is zero when the correlation of structural residuals is zero, and non-zero when the correlation of structural residuals is non-zero. See Appendix D for the full set of parameters $\phi_{12} = 0.038$, and as a result

$$0.039 = cov(\varepsilon_{it}^{a}, \varepsilon_{it}^{e}) = -1 + 2 * \frac{\exp(0.038)}{1 + \exp(0.038)}$$

D. Extension to Trivariate Ordered Logit for Employment, Alcohol and Cigarette Use

While the relationship between cigarettes and employment is beyond the scope of this study, I extend the bivariate ordered logit model into a trivariate ordered logit by adding a third equation corresponding to the latent utility for cigarettes in order to address any potential confounding effects that arise as alcohol and cigarettes may be related (Decker and Schwartz, 2000).

The levels for alcohol and employment are defined as before. I define cigarette abstention as no cigarettes in the last month, heavy smoking is defined as at least ten cigarettes in the last month, and moderate smoking is defined as those that are neither cigarette abstainers nor cigarette heavy smokers.

Table 6 presents the coefficients γ_{kj}^{Low} and γ_{kj}^{High} for k, j={alcohol, cigarette, employment}. The coefficients from Panel A Table 7 can be summarized as follows. First, all three outcomes exhibit strong state dependence *within* outcome, which increases in magnitude with dosage level. That is, $\gamma_{jj}^{Low} > 0$, $\gamma_{jj}^{High} > 0$, and $\gamma_{jj}^{High} > \gamma_{jj}^{Low}$. In particular, $\gamma_{a,a}^{Low} = 1.587$ (SE = 0.037), $\gamma_{a,a}^{High} = 4.388$ (SE = 0.173), $\gamma_{e,e}^{Low} = 0.997$ (SE = 0.037), $\gamma_{e,e}^{High} = 2.381$ (SE = 0.034), $\gamma_{c,c}^{Low} = 1.659$ (SE = 0.057), $\gamma_{c,c}^{High} = 4.227$ (SE = 0.069).

Second, regarding the coefficients associated with the state dependence *between* alcohol and employment the results are in line with the bivariate ordered logit model, as lagged moderate alcohol consumption increases future employment while heavy alcohol consumption does not have a statistically significant effect on employment ($\gamma_{a,e}^{Low} > 0$, $\gamma_{a,e}^{High} = 0$). In particular, $\gamma_{a,e}^{Low} = 0.304$ (SE = 0.035), $\gamma_{a,e}^{High} = 0.068$ (SE = 0.143). In line with the results of the bivariate ordered logit model, the coefficients do not suggest a statistically significant effect from lagged employment to future alcohol use, as, $\gamma_{e,a}^{Low} = 0.245$ (SE $= 0.051$), $\gamma_{e,a}^{High} = 0.299$ (SE =0.042).

Third, the coefficients associated with the state dependence between cigarettes and employment indicate that cigarette consumption, whether moderate or heavy, is detrimental to future employment ($\gamma_{c,e}^{Low} < 0$, $\gamma_{c,e}^{High} < 0$).). In particular, $\gamma_{c,e}^{Low} = -0.204$ (SE $= 0.054$), $\gamma_{c,e}^{High} = -0.244$ (SE $= 0.058$).

Fourth, regarding determinants of cigarette consumption, moderate alcohol consumption and full-time employment decrease the probability of cigarette consumption, as $\gamma_{a,c}^{Low} = -0.1273$ (SE $= 0.0487$), $\gamma_{e,c}^{High} = -0.1572$ (SE $= 0.0488$).

Fifth, results remain consistent as I estimate this analysis without restricting it to poorly educated individuals (Panel B), restricting it to those "likely former alcohol misuser," defined as individuals that went through that dramatic change of stopping alcohol consumption entirely at any point (Panel C) and those who went through that change at any point between the waves of interest 2002-2011 (Panel D). The results also remain consistent with an alternative definition of heavy drinker, which define them as those who consumed alcohol at least 20 days in the last month but only 70 percent of those days involve binge drinking.

Following the previous section, I compute the average partial effects (APEs) using the structural estimates in order to directly interpret the effect of alcohol and cigarettes by dosage level on employment, holding preferences constant. Given that individuals are only observed in one possible scenario ($Y_{i,c,t-1}^{None} = 1$ or $Y_{i,c,t-1}^{Low} = 1$, or $Y_{i,c,t-1}^{High} = 1$), observed data would only allow me to compute either $P[Y_{i,e,t}^{High} = 1 | Y_{i,c,t-1}^{Low} = 1]$ or $P[Y_{i,e,t}^{High} = 1 | Y_{i,c,t-1}^{High} = 1]$ or $P[Y_{i,e,t}^{High} = 1 | Y_{i,c,t-1}^{None} = 1]$. Therefore, I hold all parameters constant and artificially set lagged cigarette consumption to moderate for everyone to compute $P[Y_{i,e,t}^{High} = 1|Y_{i,c,t-1}^{Low} = 1|\alpha_i]$, and then I can artificially set lagged cigarette consumption to heavy consumption for everyone to compute $P[Y_{i,e,t}^{High} = 1|Y_{i,c,t-1}^{High} = 1|\alpha_i]$ and average it over individuals and periods.

The APE from the trivariate ordered logit models are displayed in Table 7 and can be summarized as follows.

First, the APEs are consistent with the bivariate ordered logit model, as there is evidence of strong state dependence for alcohol and employment. In addition, the APE indicates that if all respondents were to artificially consume heavy levels of cigarettes in any given period holding their preferences constant, the probability of being heavy smokers in next period would increase by 40 percentage points relative to the counterfactual where they are abstainers.[36] It is worth to note that the observed $P\left[Y_{i,c,t}^{High} = 1 \middle| Y_{i,c,t-1}^{High} = 1\right] = 0.79$ while the simulated $P[Y_{i,c,t}^{High} = 1|Y_{i,c,t-1}^{High} = 1|\alpha_i^c]$=0.425, and this large gap supports the notion that the observed persistence is highly driven by unobserved heterogeneity.

Second, the APEs are consistent with the bivariate ordered logit model, as full-time employment is higher among those with moderate alcohol consumption in the past period than among alcohol abstainers or heavy users. Third, the APEs indicate that full-time employment is higher among those who abstain from cigarettes, and any dosage level of cigarette consumption decreases the likelihood of full-time employment. In particular, the model predicts that the probability of full-time employment is highest if respondents are assigned abstention from cigarettes in the previous period ($P[Y_{i,e,t}^{High} = 1|Y_{i,c,t-1}^{None} = 1|\alpha_i]$=0.6). That probability decreases to

[36] $P[Y_{i,c,t}^{High} = 1|Y_{i,c,t-1}^{High} = 1|\alpha_i^c]$ - $P[Y_{i,c,t}^{High} = 1|Y_{i,c,t-1}^{None} = 1|\alpha_i^c]$=0.42-0.02=0.4

$P[Y^{High}_{i,e,t} = 1|Y^{High}_{i,c,t-1} = 1|\alpha_i]$ =0.56 and $P[Y^{High}_{i,e,t} = 1|Y^{Low}_{i,c,t-1} = 1|\alpha_i]$ =0.57 when artificially assigned high and moderate levels of cigarette consumption. Overall, any level of cigarette consumption is detrimental to future full-time employment.

E. <u>Policy Consequences of Moderate versus Heavy Alcohol Use</u>

Among the main attractions of structural models is that they allow the researcher to evaluate policy counterfactual scenarios. That is, using the parameter estimates and holding preferences constant, I can evaluate the extent to which artificially replacing heavy alcohol consumption with moderate consumption for all observations affects the simulated probability of full-time employment.

This subsection simulates full-time employment under several different simulated policy counterfactual scenarios. Figure 1 indicates that the bivariate ordered logit model simulates closely the full-time employment patterns over the years[37]. Using the parameters estimated by the bivariate ordered logit model, I simulate full-time employment after artificially assigning alcohol abstention, moderate alcohol consumption, and heavy alcohol consumption to all respondents (Figure 2) holding preferences constant. Surprisingly, artificially replacing predicted alcohol consumption, whether heavy or moderate, with alcohol abstention decreases full-time employment (policy simulation 1) and artificially assigning all respondents heavy alcohol consumption decreases employment even further (policy simulation 2). Finally, artificially assigning moderate alcohol consumption to everyone increases full-time employment (policy simulation 3) relative to the simulated data with alcohol abstention or heavy alcohol use.

I revisit this exercise using parameter estimates from the trivariate ordered logit model after confirming that the simulated data reflects the patterns of full-time employment in the data

[37] See coefficients in Table 8.

(Figure 3). The policy simulations presented in Figure 4 are in line with our previous results that artificially replacing alcohol outcomes with moderate alcohol consumption increases full-time employment relative to the baseline, and even more so relative to alternative policy counterfactuals where all respondents are artificially assigned either to alcohol abstention or heavy alcohol consumption. Interestingly, policy simulations indicate that cigarette consumption at any dosage level is detrimental to full-time employment. That is, abstention increases full-time employment relative to the baseline, and cigarette consumption (moderate or heavy) decreases (Table 9)

Taken together, these policy counterfactuals indicate that, holding preferences constant, moderate alcohol consumption increases full-time employment while any dosage level of cigarette consumption is detrimental to full-time employment. While determining the mechanisms behind moderate alcohol increasing full-time employment is beyond the scope of this study, a potential mechanism could be social capital accumulation that occurs during social events that involve moderate alcohol consumption.

These policy counterfactuals are particularly valuable in the absence of policies that differentially target moderate and heavy alcohol consumption, and sheds light on potential negative externalities of currently existing blunt policies that do not distinguish between dosage levels and may end up lowering employment in ways that have not previously been expected.

VI. CONCLUSION

Given the prevalence of alcohol consumption among working age individuals[38] and the importance of employment for economic prosperity, it is crucial for policy to understand the

[38] https://www.cdc.gov/alcohol/fact-sheets/binge-drinking.htm

relationship between alcohol use and individual's success in the labor market. The mixed findings regarding the relationship between alcohol and employment are unsurprising given that moderate and heavy alcohol consumption may shift employment in opposite directions. As alcohol-related policies introduce variation in the cost of alcohol access without differentially targeting moderate from heavy alcohol consumption, studies that rely on such policies (e.g. variation in alcohol taxes) estimate the aggregate effect of both moderate and heavy alcohol consumption. Due to the disproportionately high social costs driven by heavy alcohol users and the potential positive externalities driven by moderate alcohol consumption, evaluating closely the relationship between alcohol and employment separating moderate from heavy alcohol consumption is crucial for policy.

This study is the first one to address this gap in the literature and separately estimate the effect of moderate versus heavy alcohol consumption on an individual's probability of maintaining full-time employment. In particular, this study develops a multivariate *ordered* logit model that examines alcohol and employment trajectories in order to separately identify habit formation from unobserved preferences while separating the effect by dosage level. This paper is, to the best of my knowledge, the first to provide causal evidence of a previously identified stylized facts: moderate alcohol consumption leads to more employment.

This study makes three contributions. First, this study extends the consideration of state dependence and unobserved heterogeneity to an *ordered* multiproduct setting, where the outcomes have dosage levels, and outcomes are not necessarily mutually exclusive. This newly-developed multiple equation model is flexible enough to account for reverse causality by dosage level, for arbitrary correlation of alcohol and employment preferences, and the initial conditions problem caused by the fact that the entire lifetime trajectory of alcohol consumption and employment is not observed by the econometrician (e.g., Heckman 1981b). In addition, this paper implements a variant of already-existing sample-analogue generalized residuals that are used in structural estimation to assess model misspecification (Card and Hyslop, 2005; Deza, 2015) and extends it to the ordered setting.

As the second contribution, this paper uses this newly-developed methodological contribution to re-evaluate the relationship between alcohol and employment. Previous literature establishes descriptive evidence of a positive association between moderate drinking and earnings (Barrett, 2002; French and Zarkin, 1995; Hamilton and Hamilton, 1997; Heien, 1996; Bockerman et al, 2017; Lee, 2003; Zarkin at al, 1998; Bockerman et al, 2017), as well as employment (Peters and Stringham, 2006; Bockerman et al, 2017). Given the lack of natural experiments that affect only one dosage level of alcohol use, previous literature that examines the relationship between alcohol and employment using these natural experiments is unable to separately identify the causal effect of alcohol consumption on employment. This study is the first to separately identify the extent to which the observed positive correlation between moderate alcohol consumption and employment reflects a causal relationship (e.g. if heavy alcohol consumption is detrimental to health while moderate consumption increases social capital) or simply a correlational relationships (e.g. if individuals with more self-control are more likely to consume alcohol in moderation and also more likely to be employed at any given point in time regardless of their alcohol consumption). This study finds that moderate alcohol consumption increases employment, and these results remain robust to several specifications.

As the third contribution, this newly-developed model allows me to use the estimated parameters to conduct policy counterfactuals. A particularly attractive feature of structural models is that they enable researchers to use the estimated parameters to hold preferences constant and simulate the outcome of interest (e.g. full-time employment) under several policy counterfactual scenarios (e.g. artificially changing alcohol consumption to moderate levels while holding preferences constant), which enables researchers to recommend policies. These policy counterfactuals indicate that full-time employment increases in response to artificially assigning moderate alcohol consumption to respondents, holding preferences constant, relative to artificially assigning heavy alcohol consumption and even alcohol abstention.

This study sheds light to better understanding the ways in which regulations governing the use and sale of alcohol affect labor market attachments in the United States and sheds light on previously undocumented gains of moderate alcohol consumption on full-time employment. The policy counterfactuals are particularly valuable in the absence of policies that differentially target moderate and heavy alcohol consumption, and sheds light on potential negative externalities of currently existing policies that affect all consumption levels, and may end up lowering employment in ways that have not previously been expected.

REFERENCES

Altonji, J., S. Cattan and I. Ware, 2010, Identifying sibling influence on teenage substance use. National Bureau of Economics Research Working Paper 16508.

Anderson, Peter (2012) Alcohol and the workplace. Barcelone, Department of Health, Governement of Catalonia

Arcidiacono, P., A. Khwaja, and L. Ouyang, 2011, Habit persistence and teen sex: Could increased access to contraception have unintended consequences for teen pregnancies? Journal of Business and Economics Statistics.

Baldwin, Marjorie, Steven Marcus (2014) The Impact of Mental and Substance Use Disorders on Employment Transitions" Health Economics 23: 332-344

Barrett, G. (2002) The effect of alcohol consumption on earnings. Economic Record 78: 79-96

Bockerman, Petri, Ari Hyytinen, Terhi Maczulkij (2017) Alcohol Consumption and Long Term Labor Market Outcomes. Health Economics 26: 275-291

Card, D. and D. Hyslop, 2005, Estimating the effects of a time-limited earnings subsidy for welfare leavers. Econometrica 73, 1723-1770.

Charles, Kerwin and Philip DeCicca (2008) Local labor market fluctuations and health: Is there a connection and for whom?

Chesher, A. and M. Irish, 1987, Residual analysis in the grouped and censored normal linear model. Journal of Econometrics 34, 33-61.

Coate, D. (1993) Moderate drinking and coronary heart disease mortality: Evidence from NHANES I and the NHANES I follow-up. American Journal of Public Health 83: 888-890

Cook, Philip and Michael Moore (1993) This tax's for you: the case for higher beer taxes. National Tax Journal 47:559-573

Cook, Philip and Michael Moore (1999) Alcohol. NBER Working Paper 6905. Cambridge, MA: National Bureau of Economic Research.

Dave, Dhaval and Robert Kaestner (2002) Alcohol Taxes and Labor Market Outcomes. Journal of Health Economics 21(3): 357-371

Decker, Sandra and Amy, Schwartz (2000) Cigarettes and Alcohol: Substitutes or Complements? National Bureau of Economic Research Working Paper 7535

Dee, Thomas (1999) The Complementarity of Teen Smoking and Drinking. Journal of Health Economics 18: 769-793

Dee, Thomas and William Evans (2003) Teen Drinking and Educational Attainment: Evidence from Two-Sample Instrumental Variables Estimates. Journal of Labor Economics 21(1)

Deza, Monica (2015) Is there a Stepping Stone Effect in Drug Use? Separating State Dependence from Unobserved Heterogeneity Within and Across Illicit Drugs. Journal of Econometrics 184(1): 193-207

Everding, Jakob and Jan Marcus (2020) The effect of unemployment on the smoking behavior of couples 29(2): 154-170

French, Michael, Johanna Catherine Maclean, Jody Sindelar and Hai Fang (2011) The morning after: Alcohol misuse and employment problems. Applied Economics 43: pp 2705-2720

French, Michael, and Gary Zarkin (1995) Is moderate alcohol use related to wages? Evidence from four worksites. Journal of Health Economics: 319-344

Gebritz, Markus, Henry Saffer and Michael Grossman (2021) The effect of changes in alcohol tax differentials on alcohol consumption. Journal of Public Economics 204: 204520

Gentzkow, Matthew (2007) Valuing new goods in a model with Complementarities: Online Newspapers. American Economic Review 97(3)

Goel, Rajeev (2008) Unemployment, Insurance and Smoking. Applied Economics 40(2):2593-2599

Gourieroux, C, A. Monfort, E. Renault, and A. Trognon , 1987, Generalized residuals. Journal of Econometrics 34(1-2), 5-32.

Grossman, Michael, Jody Sindelar, John Mullahy, Richard Anderson (1993) Policy Watch: Alcohol and Cigarette Taxes. Journal of Economic Perspectives 7(4): 211-222

Hamilton V. and Hamilton B, (1997) Alcohol and earnings: does drinking yield a wage premium? Canadian Journal of Economics 30: 1356-151

Heckman, J., 1981a, Heterogeneity and state dependence, in S. Rosen: Studies in Labor Markets. Chicago: University of Chicago Press, 1981, pp. 91-140

Heckman, J., 1981b, The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process, in C. F. Manski and D. McFadden: Structural Analysis of Discrete Data with Econometric Applications. Cambridge, Mass.: MIT Press, 1981, pp. 179-95

Heckman, J. and B. Singer, 1984, A method for minimizing the impact of distributional assumptions in econometric models for duration data. Econometrica 52(2), 271-320.

Heien, D.M. (1996) Do drinkers earn less? Southern Economic Journal 63: 60-68

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

Johansson, E., P. Bockerman and A. Uutela, (2009) Alcohol consumption and sickness absence: Evidence from microdata. European Journal of Public Health 19: 19-22

Keane, M., 1997, Modeling heterogeneity and state dependence in consumer choice behavior. Journal of Business & Economic Statistics 15(3), 310-327.

Kenkel, D.S. (1993) Drinking, driving and deterrence. Journal of Law and Economics 36: 877-913

Kenkel, D.S. and D.C. Ribar (1994) Alcohol consumption and young adults' socioeconomic status. Brookings Papers on Economic Activity: Microeconomics: 119-175

Laaksonen M. et al, (2009) Health-related behaviors and sickness absence from work. Occupational and Environmental Medicine 66: 840-847

Lee, YL (2003) Wage Effects of drinking in Australia. Australian Economic Review 36: 265-282

Lehman, W.E.K. and D.D.Simpson (1992) Employee health and performance in the workplace: A Survey of municipal employees of a large southwest city, IBR Technical Report (Texas Christian University, Institute of Behavioral Research, Fort Worth)

Linn, S. M. Carroll, C. Johnson, R. Fulwood, W. Kalsbeek, R. Briefel (1993) High-density lipoprotein cholesterol and alcohol consumption in U.S. white and black adults: Data from NHANES II. American Journal of Public Health 83: 811-816

McCall, B., 1994, Specification diagnostics for duration models: A martingale approach. Journal of econometrics 60(1-2), 293-312.

Marmot, M. and E. Brunner (1991) Alcohol and cardiovascular disease: The status of the U-shaped curve, British Medical Journal 303: 565-568

Mullahy, J. and J.L. Sindelar, (1993) Alcoholism, work and income. Journal of Labor Economics 11: 494-520

Mullahy, John and Jody Sindelar (1996) Employment, Unemployment and Problem Drinking. Journal of Health Economics 15: pp 409-434

Norstrom, T. (2006) Per capital alcohol consumption and sickness absence. Addiction 101 : 1421-1427

Norstrom T. and IS Moan, (2009) Per capita alcohol consumption and sickness absence in Norway. European Journal of Public Health 16(4): 383-388

Peters, Bethany and Edward Stringham (2006) No Booze? You May Lose: Why Drinkers Earn More Money than Non drinkers"

Roche AM et al (2008) Worker's drinking patterns: the impact on absenteeism in the Australian workplace. Addiction, 103: 738-748

Ruhm, Christopher (1995) Economic conditions and alcohol problems. Journal of Health Economics 14: 583-603

Ruhm, Christopher (2000) Are Recessions Good for your Health? The Quarterly Journal of Economics

Salonsalmi A., M. Laaksonen, E. Lahelma, O. Rahkonen (2012) Drinking habits and disability retirement. Addiction 107: 2128-2136

Substance Abuse and Mental Health Services Administration, 2011, *Results from the 2019 National Survey on Drug Use and Health: Graphics from the Key Findings Report* https://www.samhsa.gov/data/sites/default/files/reports/rpt29393/2019NSDUHFFRBriefSlides082120.pdf

Terza, Joseph (2002) Alcohol Abuse and Employment: A Second Look. Journal of Applied Econometrics 17: 393-404

Zarkin, G. M. French, T. Mroz, J. Bray (1998) Alcohol use and wages: new results from the National Household Survey on Drug Abuse. Journal of Health Economics 17: 53-68

Appendix

Appendix A: Complete Set of Coefficients

This Online Appendix provides supplementary tables that present all estimated coefficients for the main bivariate ordered logit model with first-order state dependence (Table A1), trivariate ordered logit model with first-order state dependence that incorporates cigarettes (Table A3), bivariate ordered probit model with first-order state dependence (Table A3), and bivariate ordered logit model with second-order state dependence (Table A4). Figure A1 displays the estimated location parameters of the unobserved heterogeneity for all three types.

Appendix B: Generalized Residuals Specification Diagnostics for Ordered Models

Given that the generalized residuals must be computed at their true random effect location parameter $((r^{j}_{i,t}(\alpha^{j}_{m}))$, which is unknown to the econometrician, I build the generalized residuals at their true random effect in the following steps.

Step 1: Assign the predicted probability of the observed sequence conditional on a type

Using the estimated parameters, I hold the random effects constant and write type-p specific likelihoods. That is, I compute the probability of the observed sequence for outcome j={alcohol, employment} for each individual i under the assumption that this individual has the random effects corresponding to type p and p=1, 2, 3.[39]

$$L^{j}_{i}(Y_{ij0,\dots,}Y_{ijT}|\alpha^{j}_{p})= \prod_{j} P\left(Y^{j}_{i,2002}\middle|\alpha^{j}_{p,0}\right) * \prod_{t=2003}^{2011} \prod_{j} P\left(Y^{j}_{i,t}\middle|\overline{Y}_{i,t-1},\alpha^{j}_{p}\right)$$

In addition, I compute the respective probability of the entire observed sequence for all outcomes for each individual i with assigned random effects of type p.

[39] The random effects vector $\alpha_{p} = (\alpha^{a}_{p}, \alpha^{e}_{p}, \alpha^{a}_{p,0}, \alpha^{e}_{p,0})$ for each type p={1,2,3} is described as follows: Type 1 $(0,0,0,0)$, Type 2 $(\alpha^{a}_{2}, \alpha^{e}_{2}, \alpha^{a}_{2,0}, \alpha^{e}_{2,0})$, and Type 3 $(\alpha^{a}_{3}, \alpha^{e}_{3}, \alpha^{a}_{3,0}, \alpha^{e}_{3,0})$.

$$L_i(\alpha_p) = \prod_j P(Y_{i,2002}^j | \alpha_{p,0}^j) * \prod_{t=2003}^{2011} \prod_j P(Y_{i,t}^j | \overline{Y}_{i,t-1}, \alpha_p^j)$$

Step 2: Calculate the conditional or posterior probabilities of being each type

Using the unconditional probability of being type p[40] (π_p), the observed sequence, and Bayes rule, I estimate the conditional (posterior) probabilities and the generalized residuals for the trivariate ordered logit model as follows.

$$w_{i,p} = P(\alpha_p | \overline{Y}_i) = \frac{L_i(\alpha_p)}{\sum_{p=1}^{P} \pi_p L_i(\alpha_p)}$$

Step 3: Build Generalized Residuals

I defined the sample-analogue generalized residual for any given type p for the ordered logit model as follows:

$$r_{it}^j(\alpha_p) = \frac{Y_{i,t}^j - E[Y_{ijt} | \alpha_i^j]}{\sqrt{var(Y_{ijt})}}$$

The $E[Y_{ijt} | \alpha_i^j]$ and $var(Y_{ijt} | \alpha_i^j)$ conditional on being type p are defined as follows:

$$E[Y_{ijt}|\alpha_i^j]=0*P(Y_{ijt} = 0 | \overline{Y}_{i,t-1}, \alpha_{ij}) + 1 * P(Y_{ijt} = 1 | \overline{Y}_{i,t-1}, \alpha_{ij}) + 2 * P(Y_{ijt} = 2 | \overline{Y}_{i,t-1}, \alpha_{ij})$$

$$\begin{aligned} Var(Y_{ijt}|\alpha_i^j) = {} & P(Y_{ijt} = 0 | \overline{Y}_{i,t-1}, \alpha_{ij}) * (0 - E[Y_{ijt} | \overline{Y}_{i,t-1}, \alpha_i^j])^2 + \\ & P(Y_{ijt} = 1 | \overline{Y}_{i,t-1}, \alpha_{ij}) * (1 - E[Y_{ijt} | \overline{Y}_{i,t-1}, \alpha_i^j])^2 + \\ & P(Y_{ijt} = 2 | \overline{Y}_{i,t-1}, \alpha_{ij}) * (2 - E[Y_{ijt|} | \overline{Y}_{i,t-1}, \alpha_i^j])^2 \end{aligned}$$

The variable $Y_{i,t}^j$is the observed outcome for drug j, where $Y_{i,t}^j$={1,2,3},and the predicted probabilities $P(Y_{ijt} = 0 | \overline{Y}_{i,t-1}, \alpha_{ij})$, $P(Y_{ijt} = 1 | \overline{Y}_{i,t-1}, \alpha_{ij})$ and $P(Y_{ijt} = 2 | \overline{Y}_{i,t-1}, \alpha_{ij})$ are computed using the estimated parameters and equations 3,4,and 5, respectively.

Step 4: Diagnose $E[\mathbf{r}_{it}^j(\boldsymbol{\alpha}_i)] = \mathbf{0}$

---

[40] The unconditional probability of being type p π_p is estimated by the model to be the share of the sample that is type p without taking into account their individual observed sequence.

The following are the individual type-specific mean generalized residuals for drug j by type

$$m0_{ij}(\alpha_p) = \frac{\sum_{t=2003}^{2011} r_{it}^{j}(\alpha_p)}{9}$$

I estimate the mean generalized residual using the conditional probabilities as weights to calculate the weighted average

$$m0_{ij} = \sum_{p=1}^{3} w_{i,p} * m0_{ij}(\alpha_p)$$

Then I take the mean over people of these residuals and report them along with the standard error to test the null hypothesis that they are zero.

Step 5: Diagnose $\mathbf{E\left[r_{it}^{j}(\alpha_i), r_{it}^{j'}(\alpha_i)\right] = 0}$

The type-specific correlation of generalized residuals is estimated as follows:

$$t0_i^{j,j'}(\alpha_p) = \frac{\sum_{t=2003}^{2011} r_{it}^{j}(\alpha_p) * r_{it}^{j'}(\alpha_p)}{9}$$

I use the conditional probabilities as weights to calculate the weighted average

$$t0_i^{j,j'} = \sum_{p=1}^{3} w_{i,p} * t0_i^{j,j'}(\alpha_p)$$

Then I take the mean over people of these residuals and report them along with the standard error to test the null hypothesis that they are zero. In a similar way, I find 1st-5th order autocorrelation and cross-product and cross-period correlations.

<u>Appendix C: Trivariate Ordered Logit Model with Second-Order State Dependence</u>

Allowing higher order lagged outcomes to affect current outcomes in a model with second-order state dependence improves upon the main specification with first-order state dependence where only outcomes from the previous year affect current outcomes. Models with higher-order state dependence still rely on the assumption of no serial correlation

($E\left[\varepsilon_{ijt}, \varepsilon_{ijt-k}\right] = 0$ for k=1,2,3,4,5), but decrease the extent to which first and second-order could be a source of misspecification.

The latent utility is now a variant of equation 1, where I incorporate second-order lagged outcomes $\sum_{k=1}^{J} \gamma_{kj}^{2Lag_{Low}} Y_{i,k,t-1}^{Low} + \sum_{k=1}^{J} \gamma_{kj}^{2Lag_{Low}} Y_{i,k,t-1}^{High}$ to the original latent utility equation 1:

$$U_{ijt} = \delta_{1j}(t - t_0) + X_{it}\beta_j + \sum_{k=1}^{J} \gamma_{kj}^{Low} Y_{i,k,t-1}^{Low} + \sum_{k=1}^{J} \gamma_{kj}^{High} Y_{i,k,t-1}^{High} + \sum_{k=1}^{J} \gamma_{kj}^{2Lag_{Low}} Y_{i,k,t-1}^{Low} + \sum_{k=1}^{J} \gamma_{kj}^{2Lag_{Low}} Y_{i,k,t-1}^{High} + \varphi_j Z_{ijt} + \alpha_{ij} + \varepsilon_{ijt}$$

Because the initial conditions are now composed of outcomes in the first two years (2002-2003), and three levels for each outcome (none, low, high), the specification of the initial conditions has 9 mutually exclusive potential outcomes (0,0) (0,1) (0,2) (1,0) (1,1) (1,2), (2,0), (2,1), (2,2), where the first (second) component indicates the dosage level of any given outcome in 2002 (2003)[41]. The initial conditions are modeled with multinomial logit since the outcomes are mutually exclusive, where the latent utility for each initial conditions equation has a separate location parameter,[42] while the non-initial conditions are still a bivariate ordered logit model.

---

[41] For example, initial conditions for employment (0,1) indicate that the respondent did not work in 2002 (employment dosage level 0) and worked part-time in 2003 (employment dosage level 1). Similarly, initial conditions for alcohol (1,2) indicate that the respondent consumed moderate levels of alcohol in 2002 (alcohol dosage level 1) and consumed heavy levels of alcohol in 2003 (alcohol dosage level 2).

[42] The model estimates the specified distribution of the unobserved heterogeneity with 3 discrete points of support, where each point of support p corresponds to the following vector of unobserved heterogeneity:
$(\alpha_p^a, \alpha_p^m, \alpha_{p,0}^{a(0,1)}, \alpha_{p,0}^{a(0,2)}, \alpha_{p,0}^{a(1,0)}, \alpha_{p,0}^{a(1,1)}, \alpha_{p,0}^{a(1,2)}, \alpha_{p,0}^{a(2,0)}, \alpha_{p,0}^{a(2,1)}, \alpha_{p,0}^{a(2,2)}, \alpha_{p,0}^{e(0,1)}, \alpha_{p,0}^{e(0,2)}, \alpha_{p,0}^{e(1,0)}, \alpha_{p,0}^{e(1,1)}, \alpha_{p,0}^{e(1,2)}, \alpha_{p,0}^{e(2,0)}, \alpha_{p,0}^{e(2,1)}, \alpha_{p,0}^{e(2,2)})$

I estimate a separate utility function with a separate random effect for each of the potential 9 outcomes, which makes the multinomial logit assumption that that the transitory shocks are independent not as restrictive[43]. The utility function at the initial conditions are specified as follows:

$$U_{i,t}^{j}(0,0) = \varepsilon_{i,t}^{j(0,0)}$$

$$U_{i,t}^{j}(0,1) = \underbrace{X_{i0}\omega_{0j}^{(0,1)} + \varphi_{j,0}^{(0,1)}Z_{i,0}^{j} + \alpha_{i,0}^{j(0,1)}}_{V_{i,0}^{j(0,1)}} + \varepsilon_{i,0}^{j(0,1)}$$

$$U_{i,t}^{j}(0,2) = \underbrace{X_{i0}\omega_{0j}^{(0,2)} + \varphi_{j,0}^{(0,2)}Z_{i,0}^{j} + \alpha_{i,0}^{j(0,2)}}_{V_{i,0}^{j(0,2)}} + \varepsilon_{i,0}^{j(0,2)}$$

$$U_{i,t}^{j}(1,0) = \underbrace{X_{i0}\omega_{0j}^{(1,0)} + \varphi_{j,0}^{(1,0)}Z_{i,0}^{j} + \alpha_{i,0}^{j(1,0)}}_{V_{i,0}^{j(1,0)}} + \varepsilon_{i,0}^{j(1,0)}$$

$$U_{i,t}^{j}(1,1) = \underbrace{X_{i0}\omega_{0j}^{(1,1)} + \varphi_{j,0}^{(1,1)}Z_{i,0}^{j} + \alpha_{i,0}^{j(1,1)}}_{V_{i,0}^{j(1,1)}} + \varepsilon_{i,0}^{j(1,1)}$$

$$U_{i,t}^{j}(1,2) = \underbrace{X_{i0}\omega_{0j}^{(1,2)} + \varphi_{j,0}^{(1,2)}Z_{i,0}^{j} + \alpha_{i,0}^{j(1,2)}}_{V_{i,0}^{j(1,2)}} + \varepsilon_{i,0}^{j(1,2)}$$

$$U_{i,t}^{j}(2,0) = \underbrace{X_{i0}\omega_{0j}^{(2,0)} + \varphi_{j,0}^{(2,0)}Z_{i,0}^{j} + \alpha_{i,0}^{j(2,0)}}_{V_{i,0}^{j(2,0)}} + \varepsilon_{i,0}^{j(2,0)}$$

$$U_{i,t}^{j}(2,1) = \underbrace{X_{i0}\omega_{0j}^{(2,1)} + \varphi_{j,0}^{(2,1)}Z_{i,0}^{j} + \alpha_{i,0}^{j(2,1)}}_{V_{i,0}^{j(2,1)}} + \varepsilon_{i,0}^{j(2,1)}$$

$$U_{i,t}^{j}(2,2) = \underbrace{X_{i0}\omega_{0j}^{(2,2)} + \varphi_{j,0}^{(2,2)}Z_{i,0}^{j} + \alpha_{i,0}^{j(2,2)}}_{V_{i,0}^{j(2,2)}} + \varepsilon_{i,0}^{j(2,2)}$$

---

[43] That is, the following transitory shocks are independent $(\varepsilon_{i,0}^{j(0,0)}, \varepsilon_{i,0}^{j(0,1)}, \varepsilon_{i,0}^{j(0,2)}, \varepsilon_{i,0}^{j(1,0)}, \varepsilon_{i,0}^{j(1,1)}, \varepsilon_{i,0}^{j(1,2)}, \varepsilon_{i,0}^{j(2,0)}, \varepsilon_{i,0}^{j(2,1)}, \varepsilon_{i,0}^{j(2,2)})$.

Appendix D: Trivariate Probit Model with First-Order State Dependence

Allowing the transitory shocks to be correlated across outcomes improves upon the main specification where only the permanent unobserved heterogeneity is arbitrarily correlated across outcomes. I relax the assumption of uncorrelated contemporaneous transitory shocks across outcomes, $\mathrm{cov}(\varepsilon_{it}^{a}, \varepsilon_{it}^{e}) = 0$, as transitory shocks for employment may be correlated with transitory shocks for alcohol (e.g. living in a high density city is associated with more job opportunities and also with more alcohol consumption). In particular, the vector $\varepsilon_{i,t}^{j}$ has the following multivariate normal distribution in the bivariate ordered probit model:

$$(\varepsilon_{it}^{a}, \varepsilon_{it}^{e}) \sim N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1 & p_{12}\\ p_{12} & 1\end{pmatrix}\right)$$

Relative to the bivariate ordered logit model, the bivariate ordered probit model estimates positive structural correlation between outcome-specific transitory shocks ($\mathrm{cov}\,(\varepsilon_{it}^{a}, \varepsilon_{it}^{e}) = 0.38$) by estimating one additional parameter ϕ_{12}.[44]

Unlike the bivariate ordered logit model, the bivariate probit model jointly models the outcome for alcohol and employment in any given period $\overline{Y}_{it} = (Y_{it}^{a}, Y_{it}^{e})$[45], and numerically approximates the probabilities of any given outcome $P\left(\overline{Y}_{it} \middle| \overline{Y}_{i,t-1}, \alpha_p\right)$, as those probabilities no longer has a closed form solution[46].

---

[44] The parameter ϕ_{12} is then used to estimate $\mathrm{cov}(\varepsilon_{it}^{a}, \varepsilon_{it}^{e})$ to be between -1 and 1. The probit model estimates parameter ϕ_{12}=0.0384 (SE=0.0857). The correlation $\mathrm{cov}(\varepsilon_{it}^{a}, \varepsilon_{it}^{e})$ is by construction between -1 and 1.

$$\mathrm{cov}\,(\varepsilon_{it}^{a}, \varepsilon_{it}^{e}) = -1 + 2 * \frac{\exp\,(\phi_{12})}{1 + \exp\,(\phi_{12})}$$

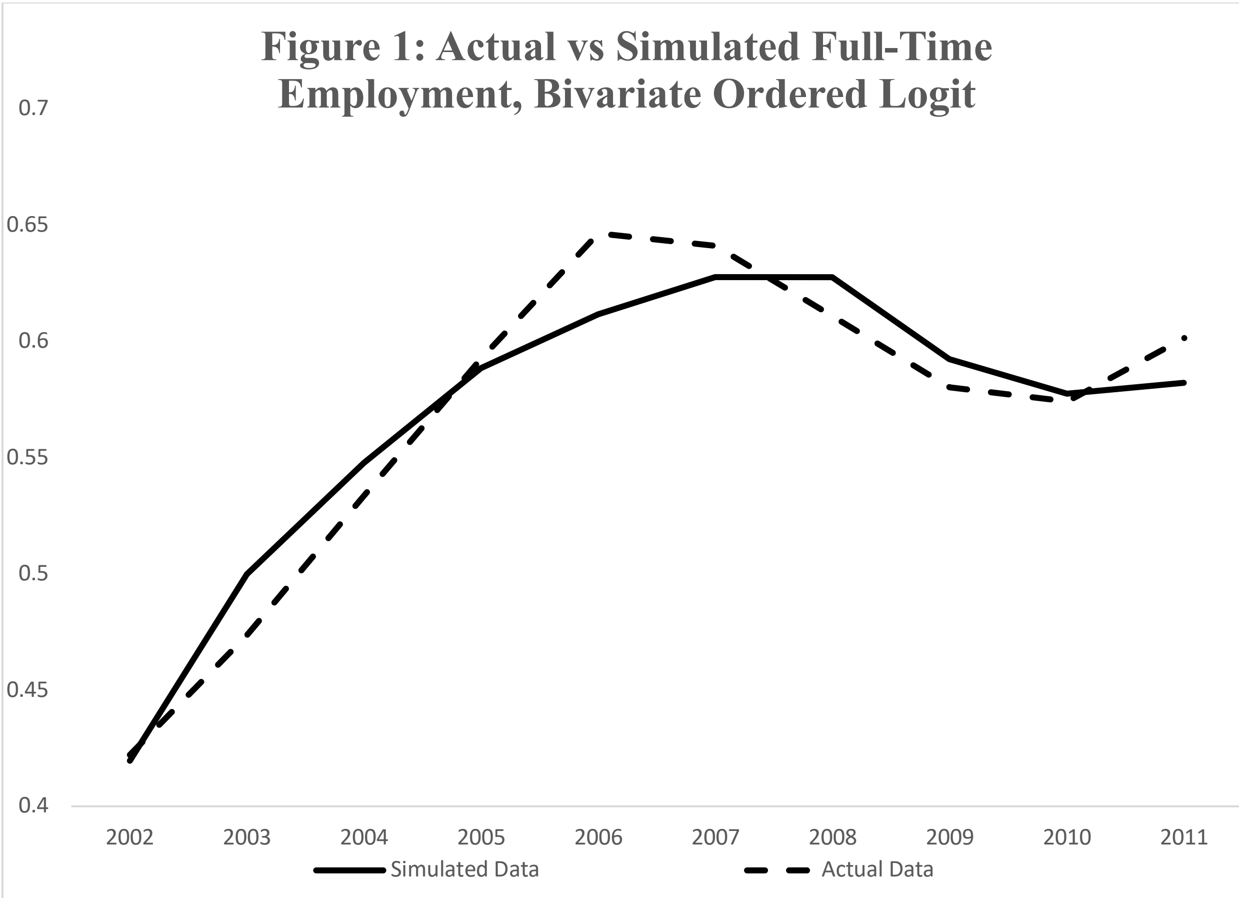
[45] In the bivariate probit model, I estimate the likelihood for the full sequence of both outcomes jointly in each period (unlike separate likelihood for each outcome in trivariate logit model) , $L_i(\overline{Y}_{i0,\dots,}\overline{Y}_{iT})$, where $L_i(\overline{Y}_{i0}, \dots . \overline{Y}_{iT}) = \sum_{p=1}^{3} \pi_p \{P(\overline{Y}_{i0}|\alpha_{p,0}) \prod_{t=2002}^{T=2011} P(\overline{Y}_{it}|\overline{Y}_{i,t-1}, \alpha_p)\}$

[46] For instance, the probability of respondent i consuming moderate level of alcohol (dosage 1) and not being employed (dosage 0) in period t is approximated as follows:

The coefficients presented in Table A3 indicate that the results are in consistent and indicate that there is state dependence within outcome which increases with dosage level, and moderate alcohol consumption increases employment.

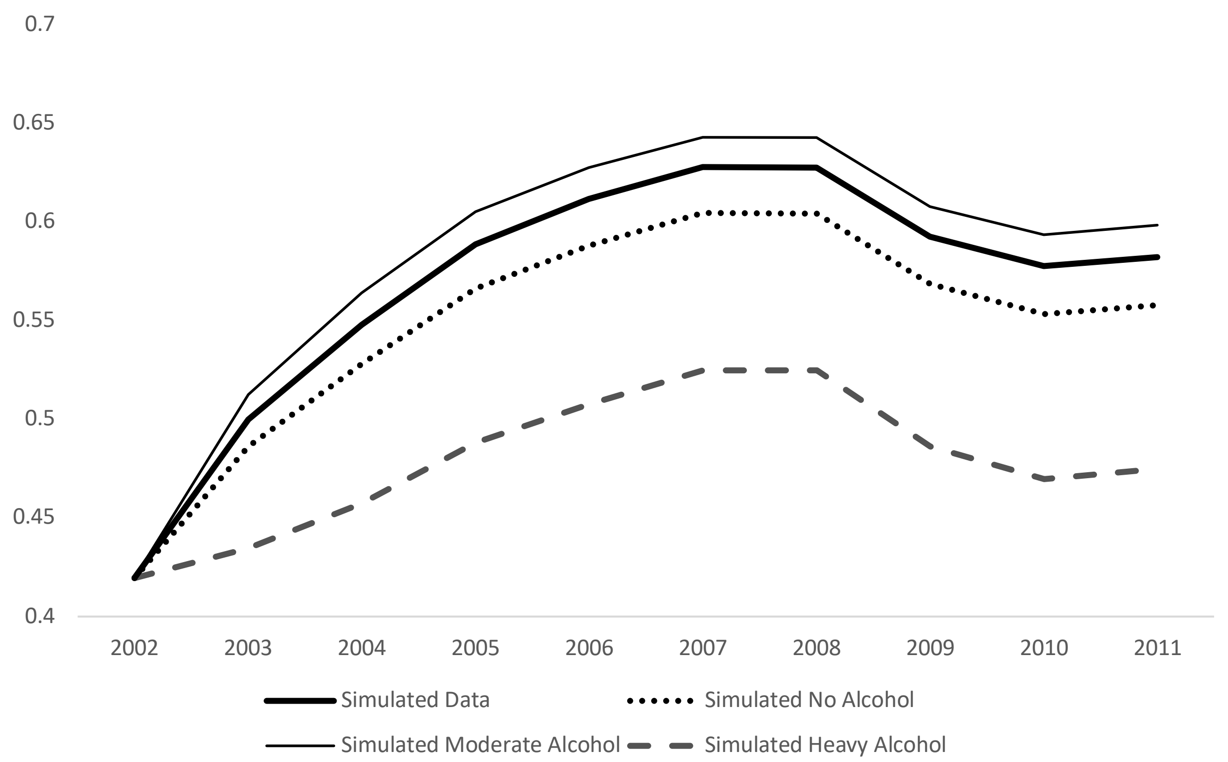
$$P\left(\overline{Y}_{it} = (1,0)\middle|\overline{Y}_{i,t-1}, \alpha_p\right) = P(\ c_{1a} < V_{it}^{a}\left(\alpha_p\right) + \varepsilon_{it}^{a} < c_{2a};\ V_{it}^{a}\left(\alpha_p\right) + \varepsilon_{it}^{a} < c_{1a})$$
$$= P(\ c_{1a} - V_{it}^{a}\left(\alpha_p\right) < +\varepsilon_{it}^{a} < c_{2a} - V_{it}^{a}\left(\alpha_p\right);\ \varepsilon_{it}^{a} < c_{1a} - V_{it}^{a}\left(\alpha_p\right))$$

Where $(\varepsilon_{it}^{a}, \varepsilon_{it}^{e}) \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & p_{12} \\ p_{12} & 1 \end{pmatrix}\right)$



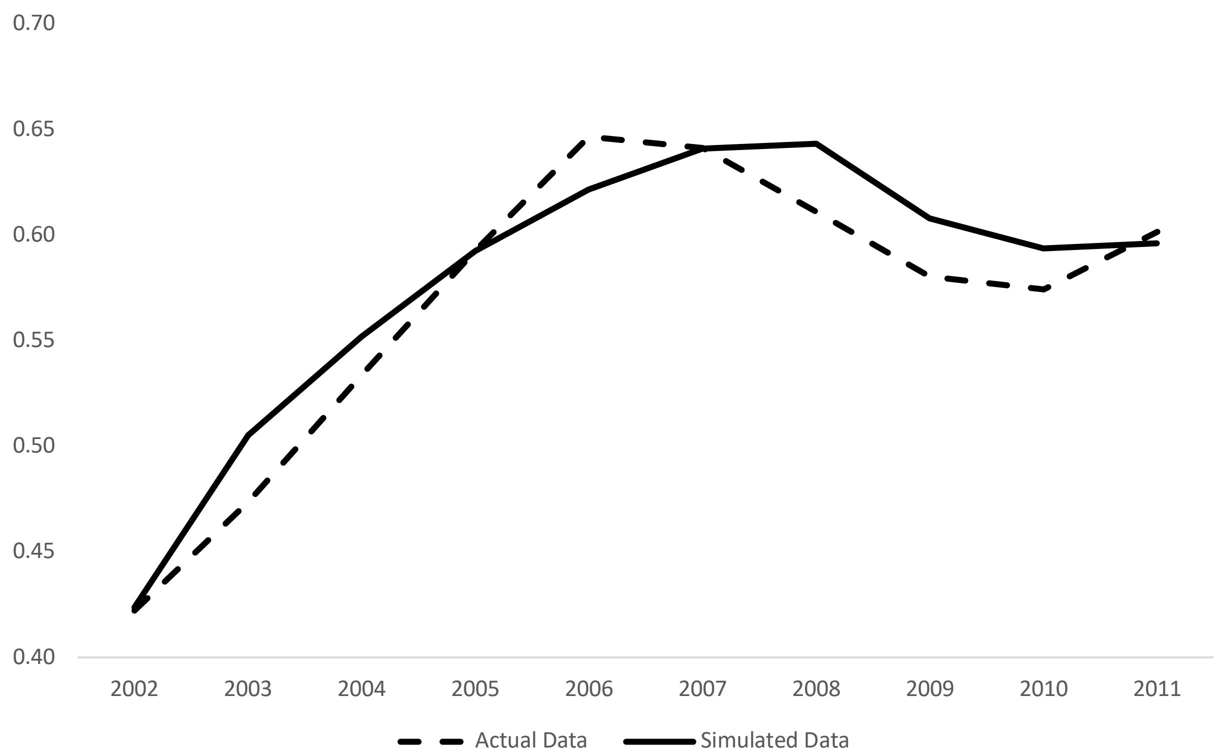
Note: This figure presents the observed and simulated probability of full-time employment using the parameters estimated by the bivariate ordered logit model.

Figure 2: Simulated Alcohol-Related Policies, Bivariate Ordered Logit



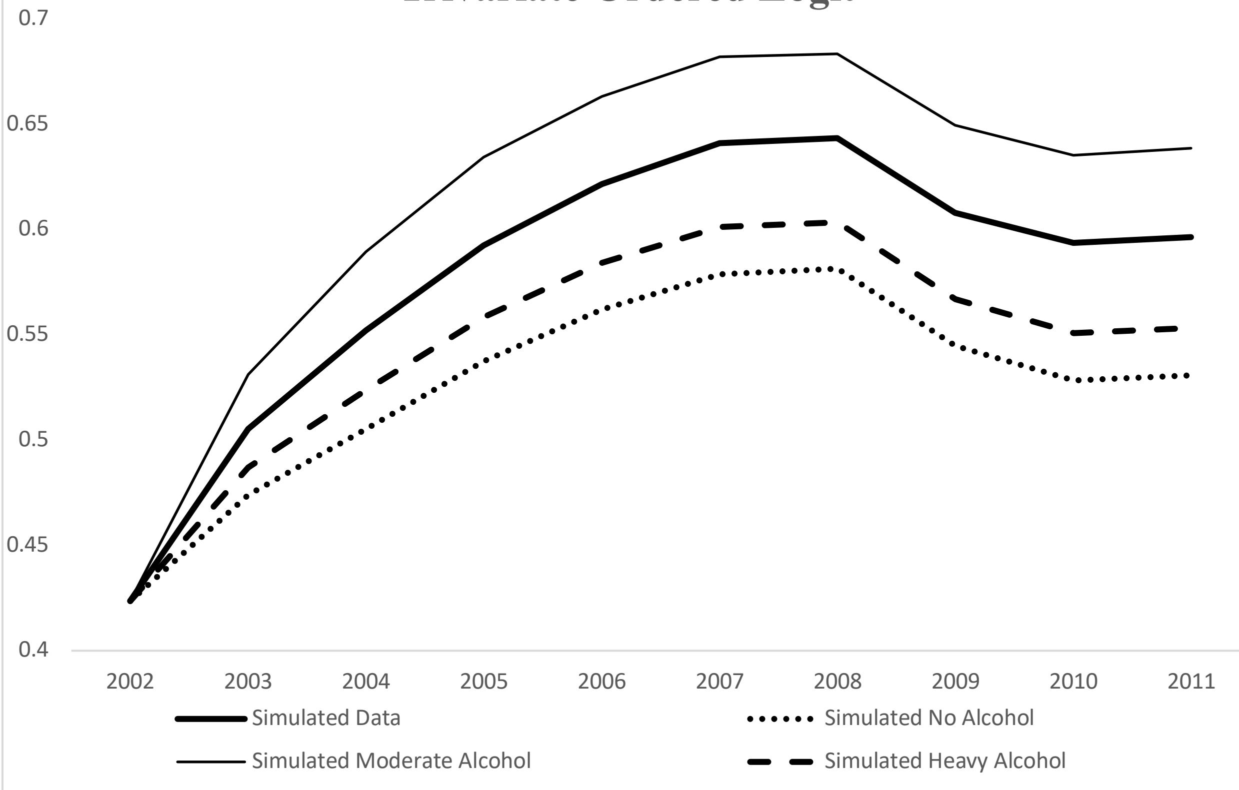
Note: This figure presents the simulated probability of full-time employment using the parameters estimated by the bivariate ordered logit model. In addition, it shoes the simulated probability of full-time employment under three counterfactual scenarios, where I hold all estimated parameters constant and artificially change the alcohol consumption dosage level for all respondents before simulating full-time employment in the next period: (i)A scenario where alcohol abstention is assigned to every respondent at every period, (ii) a scenario where moderate alcohol consumption is assigned to every respondent at every period, and (iii) a scenario where heavy alcohol consumption is assigned to every respondent at every period. See Table 8 for the respective coefficients.

Figure 3: Actual vs Simulated Full-Time Employment, Trivariate Ordered Logit



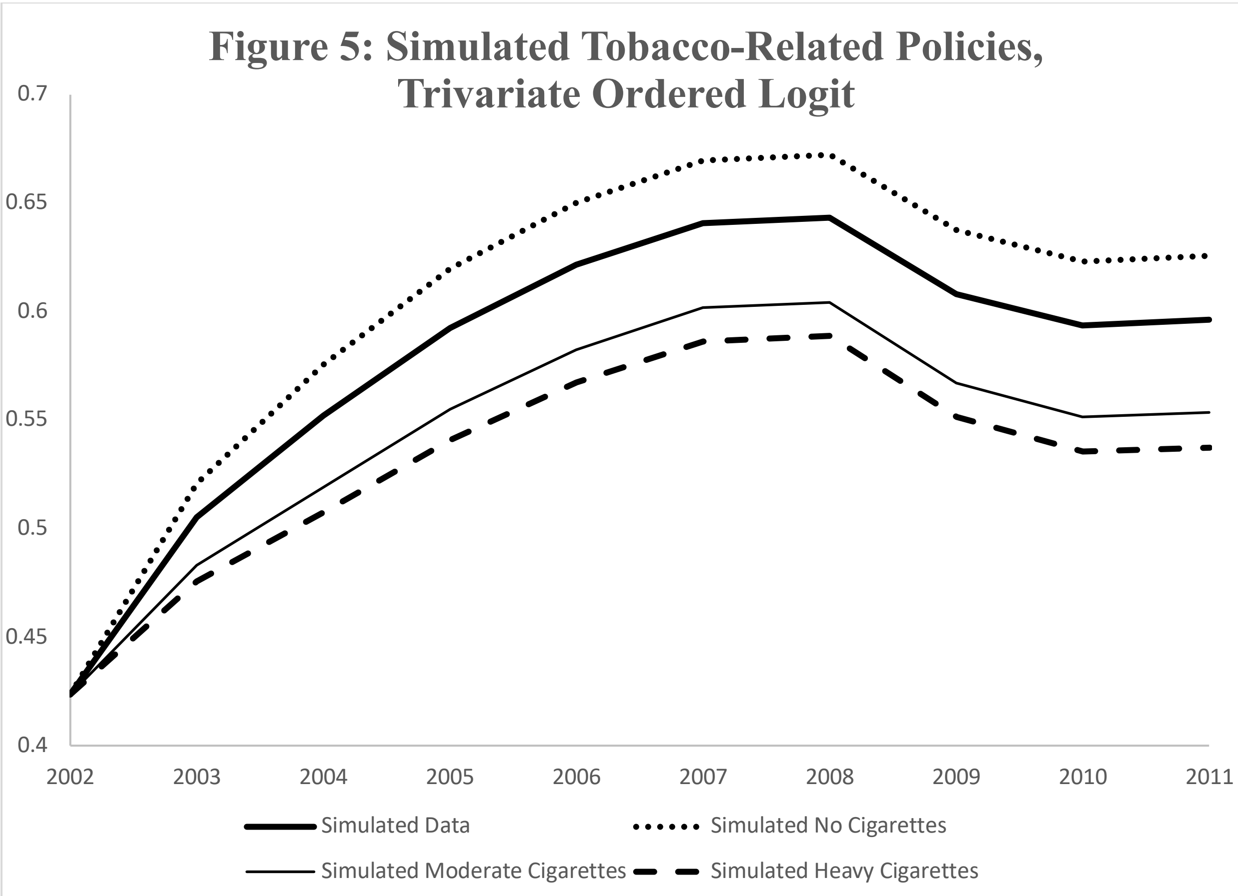
Note: This figure presents the observed and simulated probability of full-time employment using the parameters estimated by the trivariate ordered logit model.

Figure 4: Simulated Alcohol-Related Policies, Trivariate Ordered Logit



Note: This figure presents the simulated probability of full-time employment using the parameters estimated by the trivariate ordered logit model. In addition, it shoes the simulated probability of full-time employment under three counterfactual scenarios, where I hold all estimated parameters constant and artificially change the alcohol consumption dosage level for all respondents before simulating full-time employment in the next period: (i)A scenario where alcohol abstention is assigned to every respondent at every period, (ii) a scenario where moderate alcohol consumption is assigned to every respondent at every period, and (iii) a scenario where heavy alcohol consumption is assigned to every respondent at every period. See Table 9 for the respective coefficients.

Figure 5: Simulated Tobacco-Related Policies, Trivariate Ordered Logit



Note: This figure presents the simulated probability of full-time employment using the parameters estimated by the trivariate ordered logit model. In addition, it shoes the simulated probability of full-time employment under three counterfactual scenarios, where I hold all estimated parameters constant and artificially change the cigarette consumption dosage level for all respondents before simulating full-time employment in the next period: (i)A scenario where cigarette abstention is assigned to every respondent at every period, (ii) a scenario where moderate cigarettel consumption is assigned to every respondent at every period, and (iii) a scenario where heavy cigarette consumption is assigned to every respondent at every period. See Table 9 for the respective coefficients.

Table 1: Summary Statistics at First Wave

	Full sample	Subsample	
		All Education	Poorly Educated
Age 2002	19.97	19.94	19.93
	(1.42)	(1.41)	(1.42)
Male (%)	0.51	0.44	0.45
	(0.50)	(0.50)	(0.50)
Resident Father (%)	0.72	0.73	0.68
	(0.45)	(0.44)	(0.47)
P(Abstain Alcohol)	0.43	0.41	0.45
	(0.49)	(0.49)	(0.50)
P(Moderate Alcohol)	0.56	0.58	0.54
	(0.50)	(0.49)	(0.50)
P(Heavy Alcohol)	0.01	0.01	0.01
	(0.09)	(0.10)	(0.10)
P(No Employment)	0.38	0.34	0.34
	(0.48)	(0.47)	(0.47)
P(Part-time Employment)	0.29	0.31	0.24
	(0.45)	(0.46)	(0.43)
P(Full-time Employment)	0.34	0.35	0.42
	(0.47)	(0.48)	(0.49)
Highest Educ HS	0.58	0.54	0.79
	(0.49)	(0.50)	(0.41)
Highest Educ BA	0.19	0.26	0.00
	(0.39)	(0.44)	(0.00)
N	8984	4050	2781

Note: Standard errors are in parenthesis. These summary statistics are unweighted.
The first column reports demographics for the entire NLSY97 sample. The second column reports demographics for individuals with non-missing answers to the relevant questions. The third column corresponds to the subsample of column 2 with less than a college education.

Table 2: Stylized Facts

Panel A: Persistence of Full-Time (High Dosage) Employment	
P (Full-time Employment| Lag No employment)	0.265
P (Full-time Employment | Lag Part-Time Employment)	0.379
P (Full-time Employment| Lag Full-Time Employment	0.791
Panel B: Persistence of Heavy Alcohol	
P (Heavy Alcohol | Lag Abstain Alcohol)	0.001
P (Heavy Alcohol | Lag Moderate Alcohol)	0.012
P (Heavy Alcohol | Lag Heavy Alcohol	0.256
Panel C: Transition to Full-Time Employment from Alcohol	
P (Full-time Employment | Lag Abstain Alcohol)	0.501
P (Full-time Employment | Lag Moderate Alcohol)	0.636
P (Full-time Employment| Lag Heavy Alcohol	0.619
Observations	2,781

Source: Author calculations with data from the NLSY97.
Note: These probabilities are computed at the yearly level from 2003 to 2011, and the table presents the average over years. For each of the outcomes, there are three dosage levels as described in the paper: None, Low and High. For employment, these dosage levels are no employment, full-time employment and part-time employment. For alcohol, these dosage levels are abstention, moderate and heavy.

Table 3: Bivariate Ordered Logit Coefficients

	Alcohol	Employment
Panel A: Main Specification		
Lag Moderate Alcohol ($\gamma_{a,j}^{Low}$)	1.38	0.14
	(0.04)	(0.04)
Lag Part-time Employment ($\gamma_{e,j}^{Low}$)	0.03	0.67
	(0.06)	(0.04)
Lag Heavy Alcohol ($\gamma_{a,j}^{High}$)	4.05	-0.28
	(0.17)	(0.14)
Lag Full-Time Employment ($\gamma_{e,j}^{High}$)	0.06	1.97
	(0.05)	(0.04)
Observations	2,781	
Panel B: Entire Sample (Low and High Skilled)		
Lag Moderate Alcohol ($\gamma_{a,j}^{Low}$)	1.47	0.18
	(0.03)	(0.03)
Lag Part-time Employment ($\gamma_{e,j}^{Low}$)	0.03	0.77
	(0.05)	(0.03)
Lag Heavy Alcohol ($\gamma_{a,j}^{High}$)	4.31	-0.22
	(0.16)	(0.13)
Lag Full-Time Employment ($\gamma_{e,j}^{High}$)	0.04	2.23
	(0.04)	(0.03)
Observations	4050	
Panel C: Likely Alcohol Misuser		
Lag Moderate Alcohol ($\gamma_{a,j}^{Low}$)	1.32	0.14
	(0.04)	(0.04)
Lag Part-time Employment ($\gamma_{e,j}^{Low}$)	0.00	0.67
	(0.06)	(0.04)
Lag Heavy Alcohol ($\gamma_{a,j}^{High}$)	4.02	-0.29
	(0.18)	(0.15)
Lag Full-Time Employment ($\gamma_{e,j}^{High}$)	0.04	1.97
	(0.05)	(0.04)
Observations	2709	

(Continued) Table 3: Ordered Logit Coefficients

	Alcohol	Employment
Panel D: Likely Alcohol Misuser During Survey		
Lag Moderate Alcohol ($\gamma_{a,j}^{Low}$)	1.28	0.14
	(0.04)	(0.04)
Lag Part-time Employment ($\gamma_{e,j}^{Low}$)	0.00	0.69
	(0.06)	(0.04)
Lag Heavy Alcohol ($\gamma_{a,j}^{High}$)	3.99	-0.28
	(0.18)	(0.15)
Lag Full-Time Employment ($\gamma_{e,j}^{High}$)	0.04	2.00
	(0.05)	(0.04)
Observations	2654	
Panel E: Alternative Alcohol Dosage Definition		
Lag Moderate Alcohol ($\gamma_{a,j}^{Low}$)	1.37	0.14
	(0.04)	(0.04)
Lag Part-time Employment ($\gamma_{e,j}^{Low}$)	0.03	0.67
	(0.06)	(0.04)
Lag Heavy Alcohol ($\gamma_{a,j}^{High}$)	3.94	-0.17
	(0.14)	(0.12)
Lag Full-Time Employment ($\gamma_{e,j}^{High}$)	0.04	1.97
	(0.05)	(0.04)
Observations	2781	

Note: The parameters γ_{kj}^{Low} and γ_{kj}^{High} in the first column correspond to the latent utility for j=alcohol and the second column correspond to the latent utility for j=employment from equation 1, which are estimated by the trivariate ordered logit model. For example, for j=employment (second column) and k=alcohol, the respective coefficients are γ_{kj}^{Low}=0.14, and γ_{kj}^{High}=-0.28. Standard errors were computed by taking the inverse of the numerical Hessian at the estimated parameter values. For each of the outcomes, there are three dosage levels as described in the paper: None, Low and High. For employment, these dosage levels are no employment, full-time employment and part-time employment. For alcohol and cigarettes, these dosage levels are abstention, moderate and heavy.

Table 4: Average Partial Effects

	Actual	Simulated	APE
Panel A: Persistence of Full-Time Employment			
P(Full-time Employment | Lag Full-time Employment)-P(Full-time Employment | Lag No Employment)			0.411
P(Full-time Employment | Lag Full-time Employment)-P(Full-time Employment | Lag Part-time Employment)			0.268
P(Full-time Employment | Lag Part-time Employment)-P(Full-time Employment | Lag No Employment)			0.143
P(Full-time Employment |Lag Full-time Employment)	0.791	0.785	0.731
P(Full-time Employment |Lag Part-time Employment)	0.379	0.453	0.463
P(Full-time Employment | Lag No Employment)	0.265	0.234	0.320
Panel B: Persistence of Heavy Alcohol			
P(Heavy Alcohol | Lag Heavy Alcohol)-P(Heavy Alcohol | Lag No Alcohol)			0.118
P(Heavy Alcohol | Lag Heavy Alcohol)-P(Heavy Alcohol | Lag Moderate Alcohol)			0.111
P(Heavy Alcohol | Lag Moderate Alcohol)-P(Heavy Alcohol | Lag No Alcohol)			0.007
P(Heavy Alcohol | Lag Heavy Alcohol)	0.256	0.188	0.121
P(Heavy Alcohol | Lag Moderate Alcohol)	0.012	0.013	0.010
P(Heavy Alcohol |Lag Abstain Alcohol)	0.001	0.001	0.003
Panel C: Transition to Full-Time Employment			
P(Full-time Employment | Lag Heavy Alcohol)-P(Full-time Employment | Lag Abstain Alcohol)			-0.056
P(Full-time Employment | Lag Heavy Alcohol)-P(Full-time Employment | Lag Moderate Alcohol)			-0.078
P(Full-time Employment |Lag Moderate Alcohol)-P(Full-time Employment | Lag Abstain Alcohol)			0.022
P(Full-time Employment |Lag Heavy Alcohol)	0.619	0.636	0.507
P(Full-time Employment |Lag Moderate Alcohol)	0.636	0.632	0.585
P(Full-time Employment |Lag Abstain Alcohol)	0.501	0.505	0.563

Note: The first, second, and third column show the observed probability, simulated probability and the average partial effects respectively. For P(Full-time Employment | Lag No Employment), the first (second) column is the observed (simulated) probability of full-time employment among those whose observed (simulated) lagged outcome was no employment in the previous period; the third column presents the average partial effect where I use the parameters estimated by the trivariate ordered logit model for all individuals, where I artificially change the lagged employment outcome to no employment for everyone while leaving all other parameters constant. Since individuals are observed under "no employment," or "part-time employment," or "full-time employment" as mutually exclusive lagged outcomes, the probability gaps (first three rows of each panel) cannot be computed from observed data.

Table 5: Robustness Checks and Sample Analogues

	Bivariate Ordered Logit First Order		Bivariate Ordered Probit First Order		Bivariate Ordered Logit Second Order	
	Alcohol	**Employment**	**Alcohol**	**Employment**	**Alcohol**	**Employment**
Panel A: Coefficients						
Lag Moderate Alcohol ($\gamma_{a,j}^{Low}$)	1.38	0.14	1.01	0.15	1.27	0.15
	(0.04)	(0.04)	(0.08)	(0.08)	(0.04)	(0.04)
Lag Part-time Employment ($\gamma_{e,j}^{Low}$)	0.03	0.67	-0.07	0.50	0.05	0.68
	(0.06)	(0.04)	(0.12)	(0.10)	(0.06)	(0.05)
Lag2 Moderate Alcohol ($\gamma_{aj}^{2Lag_{Low}}$)	X	X	X	X	0.85	X
					(0.04)	
Lag2 Part-time Employment ($\gamma_{ej}^{2Lag_{Low}}$)	X	X	X	X	X	0.18
						(0.05)
Lag Heavy Alcohol ($\gamma_{a,j}^{High}$)	4.05	-0.28	1.40	0.07	3.54	-0.23
	(0.17)	(0.14)	(0.36)	(0.35)	(0.20)	(0.15)
Lag Full-Time Employment ($\gamma_{e,j}^{High}$)	0.06	1.97	0.08	1.09	0.02	1.85
	(0.05)	(0.04)	(0.11)	(0.09)	(0.05)	(0.04)
Lag2 Heavy Alcohol ($\gamma_{aj}^{2Lag_{High}}$)	X	X	X	X	2.23	X
					(0.21)	
Lag2 Full-Time Employment ($\gamma_{ej}^{2Lag_{High}}$)	X	X	X	X	X	0.67
						(0.04)
Observations	2,781		250		2,781	

(Continued) Table 5: Robustness Checks and Sample Analogues

	Bivariate Ordered Logit First Order		Bivariate Ordered Probit First Order		Bivariate Ordered Logit Second Order	
	Alcohol	**Employment**	**Alcohol**	**Employment**	**Alcohol**	**Employment**
Panel B: Mean, Variance, and Serial Correlation of Generalized Residuals						
$E[r_{i,t}^{j}(\alpha_i)]$	0.00	-0.01	0.02	0.00	0.00	0.00
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
$E[r_{i,t}^{j}(\alpha_i)^2]$	1.00	1.05	0.99	1.05	0.99	1.05
	(0.01)	(0.01)	(0.06)	(0.05)	(0.02)	(0.02)
$E[r_{i,t}^{j}(\alpha_i), r_{i,t-1}^{j}(\alpha_i)]$	-0.03	-0.05	-0.02	-0.02	-0.01	0.00
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
$E[r_{i,t}^{j}(\alpha_i), r_{i,t-2}^{j}(\alpha_i)]$	0.11	0.08	0.08	0.07	-0.03	-0.02
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
$E[r_{i,t}^{j}(\alpha_i), r_{i,t-3}^{j}(\alpha_i)]$	0.07	0.07	0.07	0.06	0.05	0.04
	(0.01)	(0.01)	(0.03)	(0.03)	(0.01)	(0.01)
$E[r_{i,t}^{j}(\alpha_i), r_{i,t-4}^{j}(\alpha_i)]$	0.06	0.05	0.03	0.05	0.05	0.04
	(0.01)	(0.01)	(0.03)	(0.03)	(0.01)	(0.01)
$E[r_{i,t}^{j}(\alpha_i), r_{i,t-5}^{j}(\alpha_i)]$	0.06	0.06	0.09	0.06	0.06	0.05
	(0.01)	(0.01)	(0.02)	(0.03)	(0.01)	(0.01)

(Continued) Table 5: Robustness Checks and Sample Analogues

	Bivariate Ordered Logit First Order		Bivariate Ordered Probit First Order		Bivariate Ordered Logit Second Order	
	Alcohol	Employment	Alcohol	Employment	Alcohol	Employment
Panel C: Correlation of Generalized Residuals Between Alcohol and Employment, Cross-Period						
$E\left[r^{a}_{i,t}(\alpha_i), r^{e}_{i,t}(\alpha_i)\right]$	0.03		0.04		0.03	
	(0.01)		(0.02)		(0.01)	
$E\left[r^{a}_{i,t}(\alpha_i), r^{e}_{i,t-1}(\alpha_i)\right]$	0.00		-0.01		0.01	
	(0.01)		(0.02)		(0.01)	
$E\left[r^{a}_{i,t}(\alpha_i), r^{e}_{i,t-2}(\alpha_i)\right]$	0.00		-0.02		0.00	
	(0.01)		(0.02)		(0.01)	
$E\left[r^{a}_{i,t}(\alpha_i), r^{e}_{i,t-3}(\alpha_i)\right]$	0.00		-0.02		0.00	
	(0.01)		(0.03)		(0.01)	
$E\left[r^{a}_{i,t}(\alpha_i), r^{e}_{i,t-4}(\alpha_i)\right]$	0.00		0.00		0.00	
	(0.01)		(0.03)		(0.01)	
$E\left[r^{a}_{i,t}(\alpha_i), r^{e}_{i,t-5}(\alpha_i)\right]$	0.00		-0.01		0.02	
	(0.01)		(0.03)		(0.01)	

Note: See section 4C and Appendix B for a description of the generalized residuals diagnostics. Panel A presents the parameters γ_{kj}^{Low} and γ_{kj}^{High} for all three models. The bivariate ordered logit model with second-order state dependence estimates four additional parameters $\gamma_{jj}^{2Lag_{Low}}$ and $\gamma_{jj}^{2Lag_{High}}$ for j={alcohol, employment}. Panel B presents the mean, variance, 1st-5th order autocorrelation of sample-analogue generalized residuals for alcohol and employment. Panel C presents the contemporaneous, as well as cross-period, correlation between the sample-analogue generalized residual of alcohol and employment. Standard errors reported in Panel A were computed by taking the inverse of the numerical Hessian at the estimated parameter values. Panel B reports the standard errors of the corresponding sample average. For instance, the SE of the sample average $E[r^{j}_{it}(\alpha_i)]$ is $SE = \frac{SD}{\sqrt{N}}$, where SD is the standard deviation of the original random variables $r^{j}_{it}(\alpha_i)$.

Table 6: Trivariate Ordered Logit for Alcohol, Cigarettes and Employment

	Alcohol	Employment	Cigarette
Panel A: Main Specification			
Lag Moderate Alcohol ($\gamma_{a,j}^{Low}$)	1.59	0.30	-0.13
	(0.04)	(0.04)	(0.05)
Lag Part-time Employment ($\gamma_{e,j}^{Low}$)	0.25	1.00	-0.08
	(0.05)	(0.04)	(0.06)
Lag Moderate Cigarettes ($\gamma_{c,j}^{Low}$)	-0.07	-0.20	1.66
	(0.06)	(0.05)	(0.06)
Lag Heavy Alcohol ($\gamma_{a,j}^{High}$)	4.39	0.07	0.16
	(0.17)	(0.14)	(0.18)
Lag Full-Time Employment ($\gamma_{e,j}^{High}$)	0.30	2.38	-0.16
	(0.04)	(0.03)	(0.05)
Lag Heavy Cigarettes ($\gamma_{c,j}^{High}$)	-0.16	-0.24	4.23
	(0.06)	(0.06)	(0.07)
Observations		2781	
Panel B: Entire Sample (Low and High Skilled)			
Lag Moderate Alcohol ($\gamma_{a,j}^{Low}$)	1.73	0.32	-0.18
	(0.03)	(0.03)	(0.05)
Lag Part-time Employment ($\gamma_{e,j}^{Low}$)	0.26	1.02	-0.11
	(0.04)	(0.03)	(0.05)
Lag Moderate Cigarettes ($\gamma_{c,j}^{Low}$)	-0.04	-0.15	1.74
	(0.05)	(0.04)	(0.05)
Lag Heavy Alcohol ($\gamma_{a,j}^{High}$)	4.74	0.09	0.08
	(0.16)	(0.13)	(0.15)
Lag Full-Time Employment ($\gamma_{e,j}^{High}$)	0.26	2.55	-0.15
	(0.04)	(0.03)	(0.04)
Lag Heavy Cigarettes ($\gamma_{c,j}^{High}$)	-0.28	-0.22	4.47
	(0.06)	(0.05)	(0.06)
Observations		4050	

(Continued) Table 6: Trivariate Ordered Logit for Alcohol, Cigarettes and Employment

	Alcohol	Employment	Cigarette
Panel C: Likely Alcohol Misuser			
Lag Moderate Alcohol ($\gamma_{a,j}^{Low}$)	1.54	0.30	-0.14
	(0.04)	(0.04)	(0.05)
Lag Part-time Employment ($\gamma_{e,j}^{Low}$)	0.24	0.99	-0.09
	(0.05)	(0.04)	(0.06)
Lag Moderate Cigarettes ($\gamma_{c,j}^{Low}$)	-0.08	-0.19	1.68
	(0.06)	(0.05)	(0.06)
Lag Heavy Alcohol ($\gamma_{a,j}^{High}$)	4.38	0.07	0.16
	(0.18)	(0.15)	(0.18)
Lag Full-Time Employment ($\gamma_{e,j}^{High}$)	0.30	2.36	-0.17
	(0.04)	(0.03)	(0.05)
Lag Heavy Cigarettes ($\gamma_{c,j}^{High}$)	-0.17	-0.24	4.23
	(0.07)	(0.06)	(0.07)
Observations		2709	
Panel D: Likely Alcohol Misuser During Survey			
Lag Moderate Alcohol ($\gamma_{a,j}^{Low}$)	1.53	0.30	-0.16
	(0.04)	(0.04)	(0.05)
Lag Part-time Employment ($\gamma_{e,j}^{Low}$)	0.25	0.99	-0.09
	(0.05)	(0.04)	(0.06)
Lag Moderate Cigarettes ($\gamma_{c,j}^{Low}$)	-0.12	-0.20	1.65
	(0.06)	(0.06)	(0.06)
Lag Heavy Alcohol ($\gamma_{a,j}^{High}$)	4.38	0.06	0.13
	(0.18)	(0.15)	(0.18)
Lag Full-Time Employment ($\gamma_{e,j}^{High}$)	0.29	2.39	-0.16
	(0.04)	(0.03)	(0.05)
Lag Heavy Cigarettes ($\gamma_{c,j}^{High}$)	-0.21	-0.22	4.20
	(0.07)	(0.06)	(0.07)
Observations		2654	

(Continued) Table 6: Trivariate Ordered Logit for Alcohol, Cigarettes and Employment

	Alcohol	Employment	Cigarette
Panel E: Alternative Alcohol Dosage Definition			
Lag Moderate Alcohol ($\gamma_{a,j}^{Low}$)	1.58	0.30	-0.13
	(0.04)	(0.04)	(0.05)
Lag Part-time Employment ($\gamma_{e,j}^{Low}$)	0.25	1.00	-0.08
	(0.05)	(0.04)	(0.06)
Lag Moderate Cigarettes ($\gamma_{c,j}^{Low}$)	-0.08	-0.21	1.67
	(0.06)	(0.05)	(0.06)
Lag Heavy Alcohol ($\gamma_{a,j}^{High}$)	4.25	0.20	-0.18
	(0.14)	(0.12)	(0.14)
Lag Full-Time Employment ($\gamma_{e,j}^{High}$)	0.29	2.38	-0.16
	(0.04)	(0.03)	(0.05)
Lag Heavy Cigarettes ($\gamma_{c,j}^{High}$)	-0.15	-0.25	4.24
	(0.06)	(0.06)	(0.07)
Observations		2781	

Note: These parameters are estimated by a trivariate ordered logit model. The parameters γ_{kj}^{Low} and γ_{kj}^{High} in the first, second and third column correspond to the latent utility for alcohol, employment and cigarettes from equation 1, respectively. Standard errors were computed by taking the inverse of the numerical Hessian at the estimated parameter values. For each of the outcomes, there are three dosage levels as described in the paper: None, Low and High. For employment, these dosage levels are no employment, full-time employment and part-time employment. For alcohol and cigarettes, these dosage levels are abstention, moderate and heavy.

Table 7: Average Partial Effects for Trivariate Logit Model for Alcohol, Cigarettes and Employment

Panel A: Persistence of Full-Time Employment	**Actual**	**Simulated**	**APE**
P(Full-time Employment | Lag Full-time Employment)-P(Full-time Employment | Lag No Employment)			0.516
P(Full-time Employment | Lag Full-time Employment)-P(Full-time Employment | Lag Part-time Employment)			0.300
P(Full-time Employment | Lag Part-time Employment)-P(Full-time Employment | Lag No Employment)			0.216
P(Full-time Employment |Lag Full-time Employment)	0.791	0.789	0.780
P(Full-time Employment |Lag Part-time Employment)	0.379	0.469	0.481
P(Full-time Employment | Lag No Employment)	0.265	0.239	0.264
Panel B: Persistence of Heavy Alcohol			
P(Heavy Alcohol | Lag Heavy Alcohol)-P(Heavy Alcohol | Lag Abstain Alcohol)			0.144
P(Heavy Alcohol | Lag Heavy Alcohol)-P(Heavy Alcohol |Lag Moderate Alcohol)			0.135
P(Heavy Alcohol | Lag Moderate Alcohol)-P(Heavy Alcohol | Lag Abstain Alcohol)			0.009
P(Heavy Alcohol | Lag Heavy Alcohol)	0.256	0.194	0.146
P(Heavy Alcohol | Lag Moderate Alcohol)	0.012	0.013	0.011
P(Heavy Alcohol | Lag Abstain Alcohol)	0.001	0.001	0.002
Panel C: Persistence of Heavy Cigarettes			
P(Heavy Cigarettes |Lag Heavy Cigarettes)-P(Heavy Cigarettes | Lag Abstain Cigarettes)			0.401
P(Heavy Cigarettes |Lag Heavy Cigarettes)-P(Heavy Cigarettes | Lag Moderate Cigarettes)			0.316
P(Heavy Cigarettes |Lag Moderate Cigarettes)-P(Heavy Cigarettes | Lag Abstain Cigarettes)			0.085
P(Heavy Cigarettes |Lag Heavy Cigarettes)	0.790	0.780	0.425
P(Heavy Cigarettes |Lag Moderate Cigarettes)	0.153	0.201	0.109
P(Heavy Cigarettes |Lag Abstain Cigarettes	0.016	0.010	0.024

(Continued) Table 7: Average Partial Effects for Trivariate Logit Model for Alcohol, Cigarettes and Employment

Panel D: Transition to FT Employment from Alcohol	**Actual**	**Sim**	**APE**
P(Full-time Employment |Lag Heavy Alcohol)-P(Full-time Employment |Lag Abstain Alcohol)			0.016
P(Full-time Employment |Lag High Alcohol)-P(Full-time Employment |Lag Moderate Alcohol)			-0.039
P(Full-time Employment |Lag Moderate Alcohol)-P(Full-time Employment |Lag Abstain Alcohol)			0.055
P(Full-time Employment |Lag Heavy Alcohol)	0.619	0.617	0.571
P(Full-time Employment |Lag Moderate Alcohol)	0.636	0.641	0.610
P(Full-time Employment |Lag Abstain Alcohol)	0.501	0.520	0.555
Panel E: Transition to Full-time Employment from Cigarettes			
P(Full-time Employment |Lag Heavy Cigarettes)-P(Full-time Employment |Lag Abstain smoke)			-0.042
P(Full-time Employment | Lag Heavy Cigarettes)-P(Full-time Employment |Lag Moderate Cigarettes)			-0.005
P(Full-time Employment | Lag Moderate Cigarettes)-P(Full-time Employment |Lag Abstain Cigarettes)			-0.036
P(Full-time Employment | Lag Heavy Cigarettes)	0.569	0.557	0.565
P(Full-time Employment | Lag Moderate Cigarettes)	0.572	0.568	0.570
P(Full-time Employment | Lag Abstain Cigarettes)	0.591	0.613	0.606

Note: The first, second, and third column show the observed probability, simulated probability and the average partial effects respectively. For P(Full-time Employment | Lag No Employment), the first (second) column is the observed (simulated) probability of full-time employment among those whose observed (simulated) lagged outcome was no employment in the previous period; the third column presents the average partial effect where I use the parameters estimated by the trivariate ordered logit model for all individuals, where I artificially change the lagged employment outcome to no employment for everyone while leaving all other parameters constant. Since individuals are observed under “no employment,” or “part-time employment,” or “full-time employment” as mutually exclusive lagged outcomes, the probability gaps (first three rows of each panel) cannot be computed from observed data.

Table 8: Policy Counterfactuals with Bivariate Logit, Simulated Full-Time Employment

	Actual	Simulated	Simulated Policy 1	Simulated Policy 2	Simulated Policy 3
			No Alcohol	Moderate Alcohol	Heavy Alcohol
2002	0.422	0.419	0.419	0.419	0.419
2003	0.474	0.500	0.486	0.512	0.435
2004	0.534	0.548	0.528	0.564	0.457
2005	0.593	0.588	0.566	0.605	0.488
2006	0.647	0.612	0.588	0.627	0.508
2007	0.641	0.628	0.604	0.643	0.525
2008	0.611	0.628	0.604	0.643	0.525
2009	0.580	0.592	0.568	0.608	0.486
2010	0.574	0.578	0.553	0.594	0.470
2011	0.602	0.582	0.558	0.598	0.475

Note: The first column presents the observed probability of full-time employment averaged over respondents by year. The second column shows the simulated probability of full-time employment using parameter estimates from the bivariate ordered logit model averaged over respondents by year. The last three columns show the simulated probability of full-time employment under three counterfactual scenarios, where I hold all estimated parameters constant and artificially change the alcohol consumption dosage level for all respondents before simulating full-time employment in the next period: (i)A scenario where alcohol abstention is assigned to every respondent at every period, (ii) a scenario where moderate alcohol consumption is assigned to every respondent at every period, and (iii) a scenario where heavy alcohol consumption is assigned to every respondent at every period.

Table 9: Policy Counterfactuals with Trivariate Logit, Simulated Full-Time Employment

	Actual	Simulated	Simulated Policy 1	Simulated Policy 2	Simulated Policy 3	Simulated Policy 4	Simulated Policy 5	Simulated Policy 6
			No Alcohol	Moderate Alcohol	Heavy Alcohol	No Cigarette	Moderate Cigarette	Heavy Cigarette
2002	0.422	0.424	0.424	0.424	0.424	0.424	0.424	0.424
2003	0.474	0.505	0.474	0.531	0.487	0.521	0.483	0.476
2004	0.534	0.552	0.505	0.589	0.524	0.576	0.519	0.508
2005	0.593	0.592	0.537	0.634	0.558	0.620	0.555	0.541
2006	0.647	0.622	0.562	0.663	0.584	0.650	0.582	0.567
2007	0.641	0.641	0.579	0.682	0.601	0.670	0.602	0.586
2008	0.611	0.643	0.581	0.683	0.603	0.672	0.604	0.589
2009	0.580	0.608	0.545	0.649	0.567	0.637	0.567	0.551
2010	0.574	0.593	0.528	0.635	0.551	0.623	0.552	0.535
2011	0.602	0.596	0.531	0.638	0.553	0.626	0.553	0.537

Note: The first column presents the observed probability of full-time employment averaged over respondents by year. The second column shows the simulated probability of full-time employment using parameter estimates from the trivariate ordered logit model averaged over respondents by year. Columns 3, 4 and 5 show the simulated probability of full-time employment under three counterfactual scenarios for alcohol-related simulated policies, where I hold all estimated parameters constant and artificially change the alcohol consumption dosage level for all respondents before simulating full-time employment in the next period: (i)A scenario where alcohol abstention is assigned to every respondent at every period, (ii) a scenario where moderate alcohol consumption is assigned to every respondent at every period, and (iii) a scenario where heavy alcohol consumption is assigned to every respondent at every period. Columns 6, 7 and 8 show the simulated probability of full-time employment under three counterfactual scenarios for cigarette-related simulated policies, where I hold all estimated parameters constant and artificially change the cigarette consumption dosage level for all respondents before simulating full-time employment in the next period: (i)A scenario where cigarette abstention is assigned to every respondent at every period, (ii) a scenario where moderate cigarette consumption is assigned to every respondent at every period, and (iii) a scenario where heavy cigarette consumption is assigned to every respondent at every period.

Figure A1: Arbitrary Correlation of Unobserved Heterogeneity

Source: Author’s calculations

Table A1: Full set of Coefficients, Bivariate Ordered Logit Model (First-Order State Dependence)

	Alcohol	Employment	Initial Conditions Alcohol	Initial Conditions Employment
Lower Threshold c1	2.18	-0.65		
	(0.15)	(0.11)		
Upper Threshold c2	8.37	0.40		
	(0.16)	(0.11)		
Trend	-0.01	-0.01		
	(0.01)	(0.01)		
Lag Moderate Alcohol ($\gamma_{a,j}^{Low}$)	1.3754	0.1428		
	(0.04)	(0.04)		
Lag Part-time Employment ($\gamma_{e,j}^{Low}$)	0.03	0.67		
	(0.06)	(0.04)		
Lag Heavy Alcohol ($\gamma_{a,j}^{High}$)	4.05	-0.28		
	(0.17)	(0.14)		
Lag Full-Time Employment ($\gamma_{e,j}^{High}$)	0.06	1.97		
	(0.05)	(0.04)		
Male	0.58	0.56	0.49	0.39
	(0.05)	(0.04)	(0.09)	(0.08)
Resident Father	0.17	0.17	0.26	0.28
	(0.05)	(0.04)	(0.10)	(0.08)
Lower Threshold c1 Initial Conditions			1.03	-1.64
			(0.19)	(0.30)
Upper Threshold c2 Initial Conditions			6.37	-0.53
			(0.27)	(0.30)
$\alpha_{type\ 2}^{j}$	0.91	-1.77	1.16	-1.73
	(0.13)	(0.07)	(0.21)	(0.19)
$\alpha_{type\ 3}^{j}$	2.51	0.03	2.38	0.17
	(0.07)	(0.06)	(0.15)	(0.11)
ϕ_2	-0.48			
	(0.12)			
ϕ_3	1.07			
	(0.07)			

(Continued) Table A1: Full set of Coefficients, Bivariate Ordered Logit Model (First-Order State Dependence)

	Alcohol	Employment	Initial Conditions	
			Alcohol	Employment
Age 18-20	-0.60	-0.64		
	(0.12)	(0.09)		
Age 21-23	0.02	-0.28		
	(0.09)	(0.07)		
Age 24-26	0.05	-0.13		
	(0.06)	(0.05)		
Age 18-19 Initial Conditions			-0.95	-1.10
			(0.13)	(0.11)
Age 20-21 Initial Conditions			-0.32	-0.44
			(0.13)	(0.11)
Beer Tax	-0.53		-1.42	
	(0.11)		(0.29)	
Unemp Rate		-0.07		-0.09
		(0.01)		(0.04)
Observations		2781		

Note: The parameters in the first column correspond to the latent utility for j=alcohol and the second column correspond to the latent utility for j=employment from equation 1, which are estimated by the bivariate ordered logit model. The parameters in the third and fourth column correspond to the latent utility for the initial conditions for alcohol and employment, respectively. Standard errors were computed by taking the inverse of the numerical Hessian at the estimated parameter values. For each of the outcomes, there are three dosage levels as described in the paper: None, Low and High. For employment, these dosage levels are no employment, full-time employment and part-time employment. For alcohol, these dosage levels are abstention, moderate and heavy.

Table A2: Full set of Coefficients,Trivariate Ordered Logit Model (First-Order State Dependence)

	Alcohol	Employment	Cigarettes	Initial Conditions		
				Alcohol	Employment	Cigarettes
Lower Threshold c1	0.52	-0.27	3.01			
	(0.12)	(0.10)	(0.17)			
Upper Threshold c2	6.49	0.71	5.55			
	(0.14)	(0.10)	(0.17)			
Trend	0.01	-0.01	-0.02			
	(0.01)	(0.01)	(0.02)			
Lag Moderate Alcohol ($\gamma_{a,j}^{Low}$)	1.58	0.30	-0.14			
	(0.04)	(0.04)	(0.05)			
Lag Part-time Employment ($\gamma_{e,j}^{Low}$)	0.24	1.00	-0.09			
	(0.05)	(0.04)	(0.06)			
Lag Moderate Cigarettes ($\gamma_{c,j}^{Low}$)	-0.07	-0.21	1.67			
	(0.06)	(0.05)	(0.06)			
Lag Heavy Alcohol ($\gamma_{a,j}^{High}$)	4.38	0.07	0.16			
	(0.17)	(0.14)	(0.17)			
Lag Full-Time Employment ($\gamma_{e,j}^{High}$)	0.30	2.38	-0.17			
	(0.04)	(0.03)	(0.05)			
Lag Heavy Cigarettes ($\gamma_{c,j}^{High}$)	-0.16	-0.25	4.24			
	(0.06)	(0.06)	(0.07)			
Male	0.42	0.47	0.17	0.39	0.36	-0.02
	(0.04)	(0.03)	(0.05)	(0.09)	(0.07)	(0.10)
Resident father	0.16	0.14	-0.01	0.28	0.29	0.11
	(0.04)	(0.03)	(0.05)	(0.10)	(0.08)	(0.11)

(Continued) Table A2: Full set of Coefficients, Trivariate Ordered Logit Model (First-Order State Dependence)

	Alcohol	Employment	Cigarettes	Initial Conditions		
				Alcohol	Employment	Cigarettes
Lower Threshold c1 Initial Conditions				-0.64	-1.75	0.46
				(0.16)	(0.14)	(0.38)
Upper Threshold c2 Initial Conditions				4.73	-0.71	2.25
				(0.25)	(0.14)	(0.38)
$\alpha^j_{\text{type 2}}$	0.09	-0.14	3.05	0.86	-0.19	3.63
	(0.07)	(0.06)	(0.09)	(0.10)	(0.09)	(0.16)
$\alpha^j_{\text{type 3}}$	-2.39	-0.27	-0.39	-2.32	-0.55	-0.15
	(0.07)	(0.06)	(0.18)	(0.18)	(0.11)	(0.26)
ϕ_2	0.24					
	(0.05)					
ϕ_3	-0.56					
	(0.07)					
Age 18-20	-0.27	-0.53	0.04			
	(0.11)	(0.08)	(0.13)			
Age 21-23	0.19	-0.27	0.05			
	(0.08)	(0.06)	(0.09)			
Age 24-26	0.11	-0.14	0.00			
	(0.05)	(0.04)	(0.06)			
Age 18-19 Initial Conditions				-0.87	-1.06	-0.36
				(0.13)	(0.11)	(0.14)
Age 20-21 Initial Conditions				-0.27	-0.46	-0.01
				(0.13)	(0.11)	(0.14)

(Continued) Table A2: Full set of Coefficients,Trivariate Ordered Logit Model (First-Order State Dependence)

	Alcohol	Employment	Cigarettes	Initial Conditions		
				Alcohol	Employment	Cigarettes
Beer Tax	-0.36			-1.32		
	(0.09)			(0.29)		
Unemp Rate		-0.06			-0.10	
		(0.01)			(0.01)	
Cig Tax			-0.03			-0.33
			(0.01)			(0.06)
Observations			2781			

Note: The parameters in the first column correspond to the latent utility for j=alcohol, the second column correspond to the latent utility for j=employment, and the third column correspond to the latent utility for j=cigarettes from equation 1, which are estimated by the trivariate ordered logit model. The parameters in the fourth, fifth and sixth column correspond to the latent utility for the initial conditions for alcohol, employment, and cigarettes respectively. Standard errors were computed by taking the inverse of the numerical Hessian at the estimated parameter values. For each of the outcomes, there are three dosage levels as described in the paper: None, Low and High. For employment, these dosage levels are no employment, full-time employment and part-time employment. For alcohol and cigarettes, these dosage levels are abstention, moderate and heavy.

Table A3: Full set of Coefficients, Bivariate Ordered Probit Model (First-Order State Dependence)

	Alcohol	Employment	Initial Conditions Alcohol	Initial Conditions Employment
Lower Threshold c1	2.90	-0.12		
	(0.41)	(0.29)		
Upper Threshold c2	6.22	0.54		
	(0.43)	(0.29)		
Trend	0.04	0.00		
	(0.03)	(0.03)		
Lag Moderate Alcohol ($\gamma_{a,j}^{Low}$)	1.01	0.15		
	(0.08)	(0.08)		
Lag Part-time Employment ($\gamma_{e,j}^{Low}$)	-0.07	0.50		
	(0.12)	(0.10)		
Lag Heavy Alcohol ($\gamma_{a,j}^{High}$)	1.40	0.07		
	(0.36)	(0.35)		
Lag Full-Time Employment ($\gamma_{e,j}^{High}$)	0.08	1.09		
	(0.11)	(0.09)		
Male	0.28	0.30	0.51	0.35
	(0.08)	(0.09)	(0.19)	(0.18)
Resident Father	0.13	-0.19	0.26	-0.27
	(0.08)	(0.11)	(0.22)	(0.21)
Lower Threshold c1 Initial Conditions			2.39	0.50
			(0.67)	(1.20)
Upper Threshold c2 Initial Conditions			4.95	1.36
			(0.74)	(1.20)
$\alpha_{type\ 2}^{j}$	2.03	-0.57	2.55	0.21
	(0.34)	(0.15)	(0.57)	(0.34)
$\alpha_{type\ 3}^{j}$	2.28	0.51	2.06	1.26
	(0.32)	(0.13)	(0.54)	(0.31)
ϕ_2	0.80			
	(0.30)			
ϕ_3	1.68			
	(0.25)			

(Continued) Table A3: Full set of Coefficients, Bivariate Ordered Probit Model (First-Order State Dependence)

	Alcohol	Employment	Initial Conditions Alcohol	Initial Conditions Employment
Age 18-20	0.11	-0.3256		
	0.2261	0.2323		
Age 21-23	0.3404	-0.0459		
	0.1657	0.1681		
Age 24-26	0.3412	-0.0544		
	0.1119	0.1127		
Age 18-19 Initial Conditions			-0.5824	-0.4081
			0.2673	0.2545
Age 20-21 Initial Conditions			-0.234	-0.2288
			0.2731	0.2612
Beer Tax	-0.1864		1.9916	
	0.2941		1.5551	
Unemp Rate		-0.0129		0.0949
		0.0208		0.1805
ϕ_{12}	0.0384			
	0.0857			

Note: The parameters in the first column correspond to the latent utility for j=alcohol and the second column correspond to the latent utility for j=employment from equation 1, which are estimated by the bivariate ordered probit model. The parameters in the third and fourth column correspond to the latent utility for the initial conditions for alcohol and employment, respectively. Standard errors were computed by taking the inverse of the numerical Hessian at the estimated parameter values. For each of the outcomes, there are three dosage levels as described in the paper: None, Low and High. For employment, these dosage levels are no employment, full-time employment and part-time employment. For alcohol, these dosage levels are abstention, moderate and heavy. The bivariate ordered probit model estimates an additional parameter, ϕ_{12}, relative to the bivariate ordered logit, where

$$\mathrm{cov}(\varepsilon_{it}^{a}, \varepsilon_{it}^{e}) = -1 + 2 * \frac{\exp(\phi_{12})}{1 + \exp(\phi_{12}) + \exp(\phi_{12})}$$

Table A4: Full set of Coefficients, Bivariate Ordered Logit Model (Second-Order State Dependence)

Panel A: Coefficients of non-Initial Conditions (Ordered Logit)	**Alcohol**	**Employment**
Lower Threshold c1	2.44	-0.28
	(0.15)	(0.12)
Upper Threshold c2	8.75	0.75
	(0.17)	(0.12)
Trend	-0.01	-0.03
	(0.01)	(0.01)
Lag Moderate Alcohol ($\gamma_{a,j}^{Low}$)	1.25	0.16
	(0.04)	(0.04)
Lag Part-time Employment ($\gamma_{e,j}^{Low}$)	0.04	0.67
	(0.06)	(0.05)
Lag Heavy Alcohol ($\gamma_{a,j}^{High}$)	3.48	-0.20
	(0.20)	(0.15)
Lag Full-Time Employment ($\gamma_{e,j}^{High}$)	0.03	1.84
	(0.05)	(0.04)
Lag2 Moderate Alcohol ($\gamma_{aj}^{2Lag_{Low}}$)	0.84	X
	(0.04)	
Lag2 Part-time Employment ($\gamma_{ej}^{2Lag_{Low}}$)	X	0.18
		(0.05)
Lag2 Heavy Alcohol ($\gamma_{aj}^{2Lag_{High}}$)	2.32	X
	(0.21)	
Lag2 Full-Time Employment ($\gamma_{ej}^{2Lag_{High}}$)	X	0.67
		(0.04)
Male	0.54	0.54
	(0.05)	(0.04)
Resident Father	0.14	0.15
	(0.05)	(0.04)

(Continued) Table A4: Full set of Coefficients, Bivariate Ordered Logit Model (Second-Order State Dependence)

$\alpha^{j}_{\text{type 1}}$	X	X
$\alpha^{j}_{\text{type 2}}$	0.84	-1.66
	(0.14)	(0.10)
$\alpha^{j}_{\text{type 3}}$	2.18	0.04
	(0.09)	(0.07)
ϕ_2	-0.52	
	(0.15)	
ϕ_3	1.20	
	(0.09)	
Age 18-20	-0.45	-0.43
	(0.14)	(0.11)
Age 21-23	0.06	-0.22
	(0.09)	(0.07)
Age 24-26	0.03	-0.15
	(0.06)	(0.05)
Beer Tax	-0.39	
	(0.11)	
Unemp Rate		-0.06
		(0.01)

(Continued) Table A4: Full set of Coefficients, Bivariate Ordered Logit Model (Second-Order State Dependence)

Panel B: Initial Conditions for Alcohol

	2002 None 2003 Low	**2002 None 2003 High**	**2002 Low 2003 None**	**2002 Low 2003 Low**	**2002 Low 2003 High**	**2002 High 2003 None**	**2002 High 2003 Low**	**2002 High 2003 High**
Male	0.23	0.16	0.21	0.71	2.04	-2.40	0.84	-0.20
	(0.14)	(1.06)	(0.14)	(0.12)	(0.47)	(5.09)	(0.44)	(0.91)
Resident Father	-0.03	-1.33	-0.06	0.41	0.07	-3.95	-0.37	-0.69
	(0.14)	(1.08)	(0.14)	(0.13)	(0.42)	(5.89)	(0.45)	(0.83)
$\alpha^{a}_{\text{type 1}}$	-1.48	-4.36	-1.28	-1.83	-5.45	-4.16	-5.37	-4.41
	(0.27)	(1.58)	(0.27)	(0.33)	(1.01)	(2.90)	(1.47)	(1.51)
$\alpha^{a}_{\text{type 2}}$	-0.40	-2.13	-0.53	0.21	-3.32	-2.32	-2.08	-2.10
	(0.31)	(1.27)	(0.32)	(0.28)	(0.85)	(2.31)	(0.78)	(1.15)
$\alpha^{a}_{\text{type 3}}$	0.56	-3.00	0.49	1.96	-2.36	-3.73	-1.62	-2.10
	(0.24)	(1.30)	(0.24)	(0.21)	(0.66)	(2.78)	(0.64)	(1.00)
Age 18-19 Initial Conditions	-0.51	-3.53	-0.62	-1.36	-1.75	-3.33	-1.71	-2.14
	(0.21)	(2.08)	(0.21)	(0.18)	(0.52)	(3.32)	(0.58)	(0.88)
Age 20-21 Initial Conditions	0.07	-1.21	-0.11	-0.34	-0.60	-3.12	-0.66	-3.50
	(0.20)	(1.03)	(0.20)	(0.18)	(0.48)	(3.22)	(0.54)	(1.83)
Beer Tax	-1.69	-0.62	-1.48	-2.37	-1.69	-0.44	-1.63	-0.50
	(0.42)	(2.82)	(0.43)	(0.37)	(1.27)	(6.59)	(1.42)	(2.33)
Unemp Rate	X	X	X	X	X	X	X	X

(Continued) Table A4: Full set of Coefficients, Bivariate Ordered Logit Model (Second-Order State Dependence)

Panel C: Initial Conditions for Employment	**2002 None 2003 Low**	**2002 None 2003 High**	**2002 Low 2003 None**	**2002 Low 2003 Low**	**2002 Low 2003 High**	**2002 High 2003 None**	**2002 High 2003 Low**	**2002 High 2003 High**
Male	0.14	0.45	0.22	0.32	0.27	0.21	0.27	0.85
	(0.18)	(0.17)	(0.19)	(0.16)	(0.17)	(0.18)	(0.20)	(0.13)
Resident Father	0.22	0.30	0.38	0.70	0.87	0.04	0.66	0.62
	(0.18)	(0.17)	(0.20)	(0.18)	(0.19)	(0.18)	(0.21)	(0.14)
$\alpha^{e}_{\text{type 1}}$	-0.51	-1.03	-0.72	-0.77	-0.61	0.15	-0.95	1.46
	(0.73)	(0.69)	(0.77)	(0.66)	(0.68)	(0.72)	(0.80)	(0.53)
$\alpha^{e}_{\text{type 2}}$	-1.69	-2.36	-2.07	-2.59	-2.71	-1.40	-4.24	-0.94
	(0.74)	(0.72)	(0.82)	(0.72)	(0.82)	(0.76)	(1.23)	(0.57)
$\alpha^{e}_{\text{type 3}}$	-0.33	-0.20	-0.39	-0.20	0.14	0.43	-0.24	2.22
	(0.70)	(0.66)	(0.74)	(0.64)	(0.66)	(0.70)	(0.77)	(0.51)
Age 18-19 Initial Conditions	0.31	0.08	0.70	0.31	-0.07	-0.59	-0.51	-1.42
	(0.27)	(0.26)	(0.33)	(0.25)	(0.26)	(0.26)	(0.28)	(0.19)
Age 20-21 Initial Conditions	0.14	0.62	0.40	0.29	0.20	0.13	0.01	-0.31
	(0.29)	(0.27)	(0.34)	(0.26)	(0.26)	(0.25)	(0.28)	(0.18)
Beer Tax	X	X	X	X	X	X	X	X
Unemp Rate	-0.11	-0.10	-0.20	-0.12	-0.17	-0.15	-0.10	-0.20
	(0.11)	(0.10)	(0.11)	(0.10)	(0.10)	(0.11)	(0.12)	(0.08)
Observations	500							

Note: The parameters in the first column Panel A correspond to the latent utility for j=alcohol and the second column correspond to the latent utility for j=employment from equation 1, which are estimated by the bivariate ordered logit model with second order state dependence. The parameters in Panel B and C correspond to the latent utilities for the initial conditions equations for alcohol and employment, respectively. The initial conditions equations are estimated with a multinomial logit model. Standard errors were computed by taking the inverse of the numerical Hessian at the estimated parameter values.

Table A5: Alternative Definitions of Alcohol Dosage Levels, Bivariate Ordered Logit Model (First-Order State Dependence)

	Alcohol	Employment
Panel A: Heavy Alcohol 15 Days of Drinking (70% of those Binge)		
Lag Moderate Alcohol ($\gamma_{a,j}^{Low}$)	1.355	0.143
	(0.040)	(0.038)
Lag Part-time Employment ($\gamma_{e,j}^{Low}$)	0.013	0.676
	(0.054)	(0.043)
Lag Heavy Alcohol ($\gamma_{a,j}^{High}$)	3.601	-0.132
	(0.113)	(0.101)
Lag Full-Time Employment ($\gamma_{e,j}^{High}$)	0.036	1.986
	(0.046)	(0.040)
Panel A: Heavy Alcohol 15 Days of Drinking (90% of those Binge)		
Lag Moderate Alcohol ($\gamma_{a,j}^{Low}$)	1.366	0.144
	(0.040)	(0.038)
Lag Part-time Employment ($\gamma_{e,j}^{Low}$)	0.009	0.672
	(0.055)	(0.043)
Lag Heavy Alcohol ($\gamma_{a,j}^{High}$)	3.623	-0.216
	(0.139)	(0.118)
Lag Full-Time Employment ($\gamma_{e,j}^{High}$)	0.044	1.980
	(0.046)	(0.040)

Note: The parameters γ_{kj}^{Low} and γ_{kj}^{High} in the first column correspond to the latent utility for j=alcohol and the second column correspond to the latent utility for j=employment from equation 1, which are estimated by the bivariate ordered logit model. Standard errors were computed by taking the inverse of the numerical Hessian at the estimated parameter values. For each of the outcomes, there are three dosage levels as described in the paper: None, Low and High. For employment, these dosage levels are no employment, full-time employment and part-time employment. For alcohol, these dosage levels are abstention, moderate and heavy.