NBER WORKING PAPER SERIES

SIN TAXES: DO HETEROGENEOUS RESPONSES UNDERCUT THEIR VALUE?

Padmaja Ayyagari
Partha Deb
Jason Fletcher
William T. Gallo
Jody L. Sindelar

Working Paper 15124
http://www.nber.org/papers/w15124

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2009

This work was supported by Grant Number RL1AA017542 from the National Institute on Alcohol Abuse and Alcoholism to Yale University and Grant Number R01AG027045 from the National Institute on Aging to Yale University. We gratefully acknowledge the alcohol price data that Michael French assembled and shared with us. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Sin Taxes: Do Heterogeneous Responses Undercut Their Value?
Padmaja Ayyagari, Partha Deb, Jason Fletcher, William T. Gallo, and Jody L. Sindelar
NBER Working Paper No. 15124
July 2009
JEL No. I1

ABSTRACT

This paper estimates the price elasticity of demand for alcohol using Health and Retirement Survey data. To account for unobserved heterogeneity in price responsiveness, we use finite mixture models. We recover two latent groups, one is significantly responsive to price but the other is unresponsive. Differences between these two groups can be explained in part by the behavioral factors of risk aversion, financial planning horizon, forward looking and locus of control. These results have policy implications. Only a subgroup responds significantly to price. Importantly, the unresponsive group drinks more heavily, suggesting that a higher price could fail to curb drinking by those most likely to cause negative externalities. In contrast, those least likely to impose costs on others are more responsive, thus suffering greater deadweight loss yet with less prevention of negative externalities.

Padmaja Ayyagari
Yale University
Epidemiology & Public Health
P.O. Box 208034, 60 College St.
New Haven, CT 06520-8034
padmaja.ayyagari@yale.edu

Partha Deb
Hunter College
Department of Economics
695 Park Avenue
Room 1524 West
New York, NY 10021
and NBER
partha.deb@hunter.cuny.edu

Jason Fletcher
Yale University
School of Public Health
60 College Street, #303
New Haven, CT 06510
jason.fletcher@yale.edu

William T. Gallo
Yale University
Epidemiology & Public Health
P.O. Box 208034, 60 College St.
New Haven, CT 06520-8034
william.gallo@yale.edu

Jody L. Sindelar
Yale School of Public Health
Yale University School of Medicine
60 College Street, P.O. Box 208034
New Haven, CT 06520-8034
and NBER
jody.sindelar@yale.edu

Introduction

Federal excise taxes on alcohol have not risen since 1991 and the real tax rate has been declining. However, many states are now actively considering raising taxes on alcohol to meet revenue shortfalls. The rising attention to increasing alcohol taxes may be due to the success of state and federal taxes on tobacco in boosting revenue. However, the case for a 'sin' tax to address externalities is not as strong for alcohol as it is for tobacco because most people drink responsibly, without hurting themselves or others. Another reason for taxing potentially 'tempting' goods has emerged relatively recently from the behavioral economics literature. Some individuals, especially in some situations, have trouble resisting temptations such as cigarettes or cannot resist drinking more than they had planned; this occurs even when the harms associated with succumbing ultimately outweigh the benefits of reducing the short term craving. These individuals lack the self-control to adhere to their long run preferences over tempting goods. The 'sophisticated' realize that succumbing to temptation does not maximize their long run utility, but in the short run, consumption is irresistible (Gruber and Koszegi, 2000). Taxation may serve as a pre-commitment device which helps to achieve the consumption level desired according to their longer view. (See Gruber and Mullainathan, 2002 and Hersh, 2005 for applications to tobacco.)

The price elasticity of demand for alcohol is a significant determinant of the potential welfare effects of increasing alcohol taxes[1]. Desired welfare effects may include raising revenue, reducing externalities, and acting as a precommitment device to self-control problems. If the goal is simply to raise revenue without concern for externalities, then a tax on a good with inelastic

1 Our empirical measure is price. There is evidence that alcohol taxes are at least fully passed through to the consumer (Kenkel, 2005; Young and Bielinska-Kwapsisz, 2002). Thus a tax increase results in an almost one to one price increase.

demand would minimize deadweight loss, ceteris paribus. However, if the goal is to address potential negative externalities, a tax on consumption imposing external harm is good policy only if demand is responsive to the tax. If the goal is to bolster self-control[2], the criteria are less precise, but at least those who desire a precommitment device should bear the tax and should be responsive to the tax. An optimal tax might have the greatest impact on those who impose the highest externalities (or who most want help with self-control) while having no impact on the majority of the drinkers who do not impose negative externalities (or do not seek self-control devices)[3]. This would reflect heterogeneity of demand across subpopulations.

While much of the economics literature on alcohol is aimed at youth drinking, we focus on older individuals and the potential impact of price on this population. Perhaps owing to the fact that many people drink less as they age, there is less research on older drinkers. However, older drinkers are an important set of drinkers in part due to critical interactions between alcohol and medication and also due to specific diseases being more common in older populations (Williams, 1984; Abrams and Alexopoulos, 1977). Further, although older individuals drink less than younger ones, they attain higher blood alcohol content for a given amount of alcohol consumed and, for any given level of blood alcohol, there is an intensified sensitivity to ethanol (Vestal et. al., 1977; Vogel-Sprott and Barrett, 1984). Further, balance, gait and cognition can be disturbed by alcohol consumption, and these issues are of greater concern as one ages (Williams, 1984).

[2] Behavioral economics literature on 'asymmetric paternalism' (Cameron et al, 2003) 'opitimal paternalism' (O'Donaghue and Rabin (2003) and others (Bernheim et al 2005) have opened the way for economists to consider self-control issues in discussions of optimal tax policy.

[3] We ignore the issue of regressivity.

In this paper, we estimate the price elasticity of demand for alcohol explicitly allowing for differences by complex sub-groups of individuals. We use a finite mixture model (FMM) methodology to allow the price elasticity to vary by latent groups. Specifically, FMM permits estimation of price elasticity to vary across complex, latent sub-groups that could not be identified by using simple groups such as by age, race, gender or consumption level[4]. Further, FMM allows prior and posterior characterization of the component groups. Our results suggest that important differences in price elasticity are revealed by our FMM approach and that these differences are masked in a single equation method that assumes homogeneity in response. By characterizing these two groups of individuals, we further find that some of the difference between these two latent groups can be explained by differences in risk aversion, locus of control and financial planning horizon; the differences are not only due to addiction to alcohol. These findings have critical policy applications that we discuss below.

We add to the literature in multiple dimensions. First, the topic of alcohol taxes is of current policy concern given the latest wave of proposed state level tax hikes on alcohol[5]. Second, our use of FMM allows heterogeneity across latent groups, identifying differences which would otherwise be masked by pooling unlikes. The heterogeneity that we find across

---

[4] Methods such as quantile regression allow for heterogeneity in the price elasticity, however they do not distinguish between whether the *behavior* of the individual at the xth percentile has changed or the *individual* at the xth percentile has changed.

[5] At least 24 states are considering proposals to raise alcohol taxes. In addition, states are considering easing restrictions on alcohol sales to boost tax revenues. Examples of policies under consideration include: in Utah, lawmakers are considering ending a 40-year-old law requiring consumers to get a license before drinking in a bar; Georgia, Connecticut, Indiana, Texas, Alabama, and Minnesota may end Sunday-sales bans and Alabama may allow sale of beer with 13.9% percent, instead of the current 6% volume. Accessed 6-23-09 http://www.jointogether.org/news/headlines/inthenews/2009/state-loosening-alcohol-laws.html

latent groups is relevant to policy goals. Third, we examine factors such as locus of control, financial planning horizon and risk aversion, thus contributing to the more modern behavioral economic literature on optimal taxation of addictive goods.

Background on price elasticity.

There is considerable range in estimates of the price elasticity of demand for alcohol with variation likely attributable to data type (aggregated or individual) and source (national, cross-sectional), age groups (often youths), measure of consumption, price (beer, wine or spirits, separately or an average), use of tax rate versus price, econometric methods and other factors. Wagenaar et al. (2009) conduct a formal meta-analysis and review of 1003 estimates of the price and tax elasticities of demand for alcohol from 112 studies. See also Leung and Phelps (1993) Ornstein (1980), Ornstein and Levy (1983) for earlier reviews. Grossman et al. (1998) review the evidence for youths specifically. Wagenaar et al. conclude that alcohol prices and tax rates significantly reduce consumption of alcohol. Using the simple mean effect across all of the studies, they find price elasticities for beer, wine, and spirits to be respectively: -.46, -.69, and -.80[6]. Using more-sophisticated meta-analysis techniques, including weighting the study outcomes by the precision of the estimates, they find smaller estimates of the elasticities. Specifically they find the elasticities for beer, wine and spirits to be respectively: -.17, -.30 and -.29. They estimate the overall price elasticity for alcohol to be -.44 based on studies using aggregate level data and -.03 for studies using individual level data. For heavy drinkers, the elasticity is estimated to be -.28. While these meta-estimates are extremely useful benchmarks, there is potential heterogeneity by age, race and tax versus price of alcohol that are masked. For example using data from the 1993 NHIS and self-reported number of days with five or more

[6] Estimates for price and tax elasticity were not separately delineated.

drinks, Kenkel (1996) finds the elasticity of -.5 for men and more than double this for women (-1.3). For youths, while most studies find that the initiation of use of alcohol, the amount consumed conditional on drinking and the overall drinking rate are sensitive to price, statistically significant effects are not always found (Chaloupka and Wechsler, 1996). As discussed below, Dave and Saffer (2007) find differences by age.

Two extant studies are more closely related to our current study. Using the National Health Interview Survey, Manning et al. (1995) examine whether the price elasticity of demand for alcohol varies by consumption levels. Heavy drinkers are hypothesized to be less responsive to price than light or moderate drinkers. They use a two-part model to separate the decision to drink from the quantity consumed conditional on being a drinker finding an overall price elasticity of -.80[7]. However, the price elasticity for drinks, conditional on being a drinker is insignificantly different from zero while price significantly affects the decision to drink (elasticity of -.55). Quantile regression is used to address the question of heterogeneity of elasticity across drinkers based on their level of consumption, finding that the most price responsive drinkers are the moderate drinkers. The median drinker has a price elasticity of -1.19. The lowest quantile drinker has a price elasticity of 0.55 while the heaviest two quantiles have elasticities of -0.49 and 0.12 respectively; all but the latter are significant. The gradients on both sides of the peak elasticity decline rather abruptly and the upper and lower tails are significantly different from the elasticity of the median. Manning et al. conclude that there is heterogeneity in the price elasticities and that failure to differentiate these groups could conceal important policy relevant information.

[7] Their price measure is a weighted national average across beer, wine, and spirits of the price (including taxes) per ounce of pure ethanol. They use ACCRA price data and use CPI adjustments. They use the national share of each beverage in overall alcohol consumption as weights.

The Dave and Saffer (2007) paper relates to our study along several dimensions. Specifically, they use both the PSID and the HRS to examine whether the beer tax elasticity of demand for alcohol varies by risk preference. They find that a higher risk aversion reduces alcohol consumption but that the tax elasticity of demand does not vary by a binary risk preference class. Further they find that older individuals are more tax responsive (using beer tax measures) than younger individuals (comparing age groups within the PSID). Participation elasticity is estimated to be between -.05 and -.04 for younger individuals in the PSID, and -.22 to -.11 for HRS participants over age 55. Tax elasticities, conditional on being a drinker using the average number of drinks per day as the dependent variable, are estimated to be between -.08 to -.27 depending on the specification. Chronic drinkers have a tax elasticity on this intensive margin of -.27 indicating that even heavy drinkers are at least as, if not more, price sensitive than other drinkers. This contrasts to the findings of Manning et al. above in which the moderate drinkers were found to be the most price sensitive. Both the Manning et al (1995) and Dave/Saffer (2007) studies suggest that there may be some heterogeneity in the demand for alcohol. We address this directly using the Finite Mixture Model (FMM) approach to allow heterogeneity in response across latent groups. The FMM has not been used to estimate price elasticities but offers multiple advantages as indicated above.

Data

We used data from the Health and Retirement Study (HRS), which is a nationally representative longitudinal survey of individuals over 50 years and their spouses. At baseline in 1992, HRS participants included 12,652 individuals from 7,702 households. Data were originally collected through face-to-face interviews, but later interviews were completed by telephone or mail. The HRS initially sampled persons in birth cohorts 1931 through 1941 in 1992 and then conducted follow up interviews biennially. In 1998, persons from the 1924 to 1930 cohort and the 1942 to 1947 cohort were added to the original sample. In 2004, persons from the 1948 to 1953 cohort were added to the survey. Our study takes data from both the original HRS and Version H of the data prepared by RAND.[8] The RAND HRS data is a subset of the HRS data containing cleaned versions of several variables. It was created by the RAND Center for the Study of Aging with the goal of making the data more accessible to researchers. More detail on the HRS is available elsewhere (Juster and Suzman 1995).

Our sample includes all individuals who were surveyed in 1996 through 2004. We exclude observations for which data were missing on any of the following variables: drinks per day, state and census region of residence, alcohol price for the corresponding state-year, age, race, gender, height and years of education. Our final sample consists of 71,802 observations.

Dependent Variable. The HRS asked respondents whether they ever drink any alcoholic beverages such as wine, beer or liquor. Approximately 46% of the sample answered yes to this question. If they answered yes, then they were asked how many drinks they have per day on the days that they drink. This question was not asked in the first two waves of the HRS; instead respondents were asked for a range of drinks per day in general. Because the questions are not consistent, we used data only from the third (conducted in 1996) through seventh wave (conducted in 2004) of the survey. For the 71,802 observations in our maximum sample, the average number of drinks is 0.62.

Independent Variables. We controlled for age, race, gender, height in meters, income and marital status, and years of education. To allow for nonlinear effects of age, we created the following age categories: 46 to 55, 56 to 65, 66 to 75, 76 to 85 and 86 plus. The reference

[8] The RAND HRS Data file is a longitudinal data that includes the most frequently used HRS variables. It was developed at RAND with funding from the National Institute on Aging and the Social Security Administration.

category was 45 or less. We also included binary indicators for the individual's census region of residence and also wave (year). The reference group for census region was Pacific and for year was 1996. Summary statistics are shown in table 1.

Alcohol Price. We merged the HRS data with state level data on prices on wine, beer and spirits. The prices were standardized per ounce of ethanol using standard drink size for each wine, beer and spirits and the concomitant ounces of ethanol per drink. Adjustments were made for state variation in cost of living and prices were adjusted for inflation over time.[9] In our primary results, the three prices were equally weighted to develop an average price rate measure. We take the log of this average price to obtain the variable, log average price.[10] In robustness checks, price on each beer, wine and spirits were entered separately.

Risk Aversion. To measure risk preferences the HRS asked respondents to choose among four different gambles. The first gamble was presented as follows: "Suppose you are the only income earner in the family, and you have a good job … You are given the opportunity to take a new and equally good job, with a 50-50 chance that it will double your income and a 50-50 chance that it will reduce your income by a third. Would you take the new job?" If the answer was "no," the respondent was presented with the second gamble: "Suppose the chances were a 50-50 chance that it would double your income and a 50-50 chance that it would cut your income by 20 percent. Would you still take the new job?" If the answer to the first question was "yes",

---

[9] We are grateful to Michael French and his colleagues at University of Miami for sharing their data on prices. The alcohol price data have been adjusted by the cost of living; alcohol is measured as per ounce of ethanol. ACCRA is the original source of data. State level data were calculated by averaging the figures from one or more cities within each state. The data are also available online from the National Tax foundation. We adjusted for inflation over the years.

[10] The prices have been handled in a variety of different ways in the literature. For example, Manning et al. (1995) weighted their price index using national data on the relative consumption of each beer, wine and spirits. However, these national averages are not necessarily representative of consumption patters for those over 50.

the interviewer asked: “Suppose the chances were a 50-50 chance that it would double your income and a 50-50 chance that it would cut your income by half. Would you still take the new job?” See Barsky et al. 1997 for more information on this variable and its validity.

Based on their choices, we created a risk aversion variable that took the value 1 if the individual chose the most risky gamble (50-50 chances of doubling their income or reducing it by half); 2 if they chose the job with even chances of doubling their income or reducing it by a third; 3 if they chose the job with even chances of doubling their income or reducing it by a fifth; and 4 if they chose to stay with their current job.

These questions were not asked in the 1994 and 1996 waves of the HRS and were also not asked if the interview was by proxy. Note that when we include measures of risk aversion our sample size is smaller. From 1998 onwards, only some respondents were selected to answer these questions based on their cohort, age and/or random selection[11]. We treated risk aversion as a time-invariant trait and used an approach that replaces the missing information with data from the previous wave for each individual. For persons who answered these questions in more than one wave, we took the mean of their answers.

Financial Planning Horizon. To measure planning behavior, the HRS asked respondents: “In deciding how much of your (family) income to spend or save, people are likely to think about different financial planning periods. In planning your (family’s) savings and spending, which of the time periods listed in the booklet is most important to you [and your (husband/wife/partner)]?” We created a variable that took the value 1 if they answered “next few months”; 2 if they answered “next year”; 3 if they answered “next few years”; 4 if they answered “next 5-10 years”; and 5 if they answered “longer than 10 years”.

[11] For details see: http://www.rand.org/labor/aging/dataprod/randhrsh.pdf

The sample of those who are asked the financial planning question is a subset of the full sample. These questions were not asked in the 1994 and 1996 waves of the HRS and they were not asked if the interview was by proxy. In 1998 and 2000, respondents were selected to answer this question based on a combination of their cohort and random selection. In 2002, individuals who were 65 years and older were not asked this question. We treated planning horizon as a time invariant trait and applied the same approach used for risk attitudes.

Locus of Control. Locus of control is a concept developed in psychology and refers to the extent to which individuals believe that they can, through their own efforts, affect events and outcomes in their lives. Individuals with a high internal locus of control believe that events result primarily from their own behavior and actions. In 2004, HRS asked about locus of control in a self-administered questionnaire that was left behind with some respondents upon completion of a core in-person interview. The questionnaire was administered to a random sample of non-institutionalized respondents and collected information on various psychosocial variables. One question asked individuals how often they feel that what happens to them is out of their control. Answers were coded on a 4-point scale with 1 representing “often”, 2 representing “sometimes”, 3 representing “not often” and 4 representing “never”. Individuals were also asked how much they agreed or disagreed with the statements: “at home, I feel I have control over what happens in most situations” and “I feel that what happens in life is often determined by factors beyond control”. Answers were coded on a 6-point scale, with 1 representing ”strongly agree” and 6 representing “strongly disagree”.

We used principle component analysis to reduce these three variables to a single factor that was increasing in the level of internal control. The factor loaded positively on the first and third variables and negatively on the second variable. Since these variables were asked only in

2004, we treated the factor as a time invariant variable and attributed the same value to all previous waves for each person. This is consistent with the view in psychology that locus of control is a personality trait.

Forward Looking. The 2004 leave-behind questionnaire also asked individuals how much they agreed or disagreed with the following statements: “I live life one day at a time and don’t really think about the future”, “I enjoy making plans for the future and working to make them a reality” and “some people wander aimlessly through life, but I am not one of them”. Answers were coded on a 6-point scale with 1 representing “strongly agree” and 6 representing “strongly disagree”. We recoded the latter two variables so that 1 represented “strongly disagree” and 6 represented “strongly agree”. Next, we used principle component analysis to create a single factor that loaded positively on all three variables. A higher value of this factor represents a more future oriented attitude[12]. For this factor too, we attributed the same value to previous waves for each individual.

CESD Score. The RAND-HRS data included an abridged version of the Center for Epidemiologic Studies-Depression (CESD) Scale (Radloff 1977). This validated score is based on eight questions that asked the respondent if during the past week they: felt depressed, felt that everything did was an effort, experienced restless sleep, could not get going, felt lonely, felt sad much of the time, enjoyed life and was happy. The variable is a count of the number of symptoms of depression.

Binge Drinking. The HRS asked respondents the number of days that they had four or more drinks on one occasion, in the last three months.

Econometric Methods

[12] The factor loadings for locus of control and forward looking attitudes are not presented here but are available upon request.

We estimate the basic econometric model for number of drinks per day (DRINKS), an integer valued variable, using the following approach:

$$E(DRINKS_i \mid PRICE_i, \mathbf{X}_i) = \exp(\alpha PRICE_i + \mathbf{X}_i'\beta) \qquad (1)$$

Where, $PRICE_i$ is the logarithm of the average alcohol price corresponding to observation i. Socioeconomic characteristics are denoted by **X**. Equation (1) is first estimated by Poisson regression. In the Poisson model, the coefficient α is the price elasticity of alcohol consumption. However, if DRINKS is drawn from distinct subpopulations, the Poisson estimate of α is the average of the effects across subpopulations and may hide substantive differences in α across the subpopulations. Thus, we also estimate equation (1) using a finite mixture model with Poisson-distributed subpopulations.

In the finite mixture model, the random variable y is postulated as a draw from a population which is an additive mixture of C distinct classes or subpopulations in proportions π_j, such that

$$g(y_i \mid \boldsymbol{\theta}, \boldsymbol{\pi}) = \sum_{j=1}^{C} \pi_j f_j(y_i \mid \boldsymbol{\theta}_j), \;\; 0 \le \pi_j \le 1, \;\; \sum_{j=1}^{C} \pi_j = 1. \qquad (2)$$

Where, the j^{th} density is $f_j(y_i \mid \boldsymbol{\theta}_j), j = 1, ..., C$ and $\boldsymbol{\theta}_j$ is the associated set of parameters. Equation (2) assumes that the proportions π_j are constant across observations.

The mixture density in the Poisson mixture for DRINKS is given by:

$$f_j(y_i \mid \boldsymbol{\theta}_j) = \frac{\lambda_{ji}^{y_i} \exp(-\lambda_{ji})}{y_i!} \qquad (3)$$

Where, $\lambda_{ji} = \exp(\alpha_j PRICE_i + \mathbf{X}_i'\beta_j)$. Other applications of mixture models include Morduch and Stern (1997) and Conway and Deb (2002) who used a mixture of normal densities, while an early application of finite mixture of Poisson densities is Wang, Cockburn and Puterman (1998).

The likelihood function corresponding to the constant probability model is:

$$L(\boldsymbol{\theta}, \boldsymbol{\pi}) = \prod_i \sum_j \pi_j f_j(y_i \mid \boldsymbol{\theta}) = \prod_i g(y_i \mid \boldsymbol{\theta}, \boldsymbol{\pi}) \qquad (4)$$

The finite mixture models are estimated using maximum likelihood and cluster-corrected robust standard errors are used throughout for inference purposes. These are implemented using the Stata package *fmm*. Starting from the initial estimates of component proportions π_j, we re-estimate the finite mixture model assuming a (prior) component probability of the form:

$$\pi_j(\mathbf{Z}_i \mid \delta) = \mathbf{Z}_i'\delta, \quad 0 \le \pi_j \le 1, \quad \sum_{j=1}^{C} \pi_j = 1. \qquad (5)$$

The prior component probability now depends on observables $\mathbf{Z}$ and so varies across observations. This model assumes that individuals with varying observable characteristics might have different probabilities of belonging to either component. The likelihood function corresponding to the variant probability model is:

$$L(\boldsymbol{\theta}, \delta) = \prod_i \sum_j \pi_j(\mathbf{Z}_i \mid \delta) f_j(y_i \mid \boldsymbol{\theta}) \qquad (6)$$

The finite mixture model provides a natural and intuitively attractive representation of heterogeneity in a finite, usually small, number of finite mixtures latent classes, each of which may be regarded as a ‘type’, or a ‘group’. The results of two studies (Heckman and Singer 1984; Laird 1978) suggest that estimates of such finite mixture models may provide good numerical approximations even if the underlying mixing distribution is continuous. In addition, the finite mixture approach is semiparametric—it does not require any distributional assumptions for the mixing variable—and under suitable regularity conditions is the semiparametric maximum likelihood estimator of the unknown density (Lindsay 1995). Econometric applications of finite mixture models include the seminal work of Heckman and Singer (1984) to labor economics, Wedel, et al. (1993) to marketing data, El-Gamal and Grether (1995) to data from experiments in decision making under uncertainty, and Deb and Trivedi (1997) to the economics of healthcare. Finite mixture models have received increasing attention in the statistics literature as well (see McLachlan and Peel, 2000, and Lindsay, 1995, for numerous applications).

A finite mixture characterization is especially attractive if the mixture components have a natural interpretation. However, this is not essential. A finite mixture may be simply a way of flexibly and parsimoniously modeling the data, with each mixture component providing a local approximation to some part of the true distribution. A caveat to the foregoing discussion is that the finite mixture model may fit the data better simply because outliers, influential observations or contaminated observations are present in the data. The finite mixture model will capture this phenomenon through additional mixture components. Hence it is desirable that such models be supported both by a priori reasoning and by meaningful *a posteriori* differences in the behavior of the latent classes.

We use our finite mixture parameter estimates to calculate the posterior probability of being in each of the latent classes. We use Bayes' Theorem to calculate the posterior probability of membership in each class, conditional on all observed covariates and outcome, as

$$\Pr(y_i \in k \mid \boldsymbol{\theta}, y_i) = \frac{f_k(y_i \mid \boldsymbol{\theta}_k)}{\sum_{j=1}^{C} \pi_j(\mathbf{Z}_i \mid \delta) f_j(y_i \mid \boldsymbol{\theta}_j)}, \forall k = 1, 2, ..., C. \tag{7}$$

Thus the posterior probability varies across observations. We use the estimated posterior probabilities to explore the determinants of class membership.

Two potentially attractive alternative econometric strategies are worth discussing, the use of quantile regression methods, and dealing with the large fraction of zeros using two part models. Quantile regressions have been used in similar contexts to study heterogeneous responses to treatments but they have two limitations vis-à-vis finite mixture models. First, quantile regressions are not well behaved in the context of count data. Second, although quantile regression methods may detect heterogeneous responses, they provide no way to characterize the source of the heterogeneity.

Two part models are ubiquitous in the health economics literature to deal with potential heterogeneity between users and non-users when the distribution of the outcome includes a substantial fraction of zeros. Although our data do include a substantial fraction of zeros, the two-part model is less attractive than the finite mixture for two reasons. First, the two-part model may be thought of as a special case of the finite mixture model in which one of the components has a degenerate distribution at zero. Second, to the extent that some individuals are occasional drinkers who go back and forth between no drinks and light drinking, the distinction between use and non-use may be less attractive than the distinction between low and high use, which is the distinction that the finite mixture model makes.

Results

Using our full sample to estimate a single Poisson regression, we find that price has a significant ($p<0.001$) impact on the number of drinks consumed with a coefficient of -0.286

(Table 2, column 1). This estimate falls within the range of estimates from Wagenaar et al. When we estimate the FMM allowing for two latent components, we find that for the first and largest group (75.4% of the sample), price also has a significant impact on the number of drinks consumed ($p < 0.001$) with a coefficient of -1.600 (Table 2, column 2). In contrast, for the second, smaller group (24.6% of the sample), price is not significant and the coefficient is -0.035 (Table 2, column 3). The differences between these two groups in both the significance and magnitude of impact are striking. This evidence suggests that the single component model is composed of two unlike groups. The larger group (Component 1) consumes 0.129 drinks per day on average, while the smaller group (Component 2) has an average consumption of 1.879 drinks per day. This heterogeneity is masked in the single equation estimation approach. The prior probability equation (Table 2, column 4) indicates that individuals who are younger, more educated, male and white are more likely to belong to the latent Component 2.

Other Independent Variables. Education is significant in all regressions in Table 2 ($p<0.001$) and increases the number of drinks in the single Poisson equation and in the sub-population, Component 1. Further, those with higher education are less likely to be in Component 1, suggesting that education may be positively correlated with the number of drinks. However, the impact of education is, perhaps surprisingly, negative for those in Component 2. These results suggest that education is associated with more drinks in general and for those in the group with the lower level of drinking, and education is associated with an increased likelihood of being in the heavy drinking group. However, for those in this higher drinking group, the impact of education is to reduce the number of drinks. The FMM results generally support previous studies suggesting that for adults, education increases drinking on average, but these results suggest a more complicated picture. More educated individuals may be more likely to drink, but less likely to drink to overindulge.

Each of the set of the age dummies is negative and significant in the single equation Poisson estimate, with the impact increasing with higher ages (the youngest age category is omitted). Similarly, for those in Component 2, as one ages, the number of drinks declines. Also, the probability of being in Component 1 increases with age. For those in Component 1, however, age is insignificant. Thus the FMM results indicate a more complicated role for age as compared to the single equation results that suggest simply that the number of drinks decline with age.

These results also indicate an interesting role for gender in affecting the number of drinks. As would be expected, males are found to be significantly more likely to drink in the single equation regression and this same relationship is found for those in Component 2. Also, males are significantly less likely to be in Component 1. However, gender is insignificant for those in Component 1, as seen in column 2. Blacks and other races are found to drink fewer drinks in the single equation regression as well as in Components 1 and 2 and more likely to be in Component 1.

Posterior Probability. In the prior probability equation of Table 2, we use a basic set of socio-economic variables (e.g. age, race and gender) to classify observations into the latent subgroups. Using the established groups, we then use a richer set of variables to characterize the two latent components. Our descriptive analysis of posterior probabilities (shown in Table 3) aims to identify factors that are correlated with group membership. Some of the variables that are of interest to us due to behavioral economic concepts (e.g. locus of control) are available only in much smaller subsets of the HRS.

Drawing on some of the literature in behavioral economics, we hypothesized that locus of control, risk-aversion, length of planning horizon and forward looking attitudes would affect the likelihood of individuals' sorting into each component. We find that higher risk aversion is significantly and positively associated with an increased likelihood of being in Component 1 as expected. Intuitively, those who are more risk averse tend to avoid the risky behavior of heavy drinking and associated harmful behaviors and outcomes (e.g. drunk driving resulting in traffic crashes and impaired balance resulting in falls). Longer financial planning horizons are significantly associated with a reduced probability of being in Component 1. Those with a long-term perspective are less affected by current price variations; their decisions are apparently based on a longer time horizon. Similarly, those with forward looking attitudes are less likely to be in Component 1. Rather surprising to us, locus of control is associated with a decreased likelihood of being in Component 2. Locus of control is a measure of 'internal' control; it could be that people naively believe that they can control their drinking but then do not, or could be that they are in control. Locus of control is not a good measure of self-control per se, but such a measure is not available in the data.

We find that depression increases the likelihood of being in Component 1. This finding is consistent with less social interaction due to depression and in turn a lower drinking rate. This finding is also consistent with the finding in Saffer and Dave, 2005, indicating that individuals with mental disorders are more sensitive to tax. We include a measure of binge drinking (the number of days in the last three months that the individual had four or more drinks on one occasion) to determine whether the components are synonymous with binge drinking. While binge drinkers are significantly less likely to reside in Component 1, the inclusion of this indicator has little impact on the significance and sign of the other independent variables. Note that the sample size in the posterior specification varies due to limits in the number of observations available for each of these less standard variables.

Robustness Checks. We conduct several different types of robustness checks. We find qualitatively similar results across the alternative specifications. In Table 4, we compare across specifications using alternative sets of covariates. In the first three columns, we display the results of the two component FMM estimates including risk aversion and financial planning horizon directly in the component price elasticity regressions rather than only in the posterior probability. Note that the sample size when these variables are included is substantially reduced (32,989 instead of the larger 71,802), as these questions were not asked of the full population. Locus of control and forward looking attitudes are not included in the FMM regressions of drinks due to the much smaller sample size.

When planning horizon and risk aversion are entered directly into the DRINKS equation, the qualitative results still hold. Price elasticity is significant and fairly large (-1.464) in Component 1, while it is insignificant (and small) in the second. The price elasticity of Component 1 is similar to our previous estimates, but we would not expect it to be exactly the same due to the smaller sample. Risk aversion is negative and significant in Component 1, indicating that those who are most risk-averse drink fewer drinks, cet. par. This mirrors the qualitative result in Dave and Saffer (2007). Risk aversion is insignificant in Component 2. Financial planning horizon is significant in both components; however, it is positive in Component 1 but negative in Component 2. These qualitative results of opposing impacts are similar to those found for education, in that a longer financial planning horizon makes one less likely to be in Component 1 (from Table 3), but more likely to drink more conditional on being in this component and less likely to drink more for those in Component 2.

The second set of FMM results in Table 4 includes state level drunk driving laws. We include these laws to test for the impact of state sentiment.[13] Inclusion of these laws helps to address the possibility that differences in tax rates and levels of alcohol consumption are determined by differences across state in sentiment toward alcohol consumption, which could bias the results. Because zero tolerance laws affect only those under age 18, this may be the purest control, as it could not have a direct impact on alcohol consumption by older individuals. Estimates of the price elasticity are changed very little by inclusion of these proxies for state sentiment.

The next set of checks examines the number of components. We initially investigated a FMM specification with two heterogeneous groups. Then we tested to determine whether one, two or three components would better represent the underlying relationships. The test criteria for each model are reported in Table 5 while the resulting price elasticities can be seen in Table 6. We use several criteria in selecting the best model: AIC/BIC and appropriate mechanism of the estimation approach. The AIC/BIC reveal that the one component Poisson regression is inferior to a specification with multiple components. This is consistent with our finding that there are two groups with importantly different price elasticities. As a comparison, we allow three groups in the FMM specification. The three component model did not converge.

In an alternative two-component approach, rather than using only latent groups we deleted all the individuals who reported no drinking over the full set of waves in which they were in the data[14]. Then we estimated a two-component FMM model. We deleted those who do not drink over all waves on the grounds that they were dedicated non-drinkers and including them in

[13] Kenkel (1996) also uses state laws related to alcohol availability and drunk driving countermeasures as explanatory variables for alcohol demand.

[14] The HRS asked respondents whether they ever drink alcoholic beverages. Unfortunately, it is not clear whether this question refers to consumption over the individual's lifetime or current consumption. Further, it is well known that individuals tend to under-report alcohol consumption in such surveys. Thus, we defined non-drinkers as those who answered no in every survey interview.

the regressions would mask the true price elasticity for light drinkers. Otherwise, light drinkers and non-drinkers would likely be included together in the two-component model. In the preferred two-component FMM specification found in Table 2, rather than deleting the 'dedicated non-drinkers' we allowed the estimation approach to determine the composition of the latent classes. Non-drinkers are considered to be at risk of becoming drinkers.

The price elasticities associated with the different model specifications are displayed in Table 6. The qualitative conclusions that we draw from all specifications of the FMM models are similar. We consistently find that, 1) price is significant for Component 1; 2) the price elasticity is larger in this component as compared to the single equation Poisson estimate and 3) price is insignificant in the other component(s). Further, the first component is the largest group, and it comprises about 75% of the population except when we delete the nondrinkers. Thus the qualitative results from our preferred model are robust to alternative specifications. Estimated price elasticities range from -0.926 to -1.600 for the full sample. Even the lowest elasticity is more than double that of the Poisson regression (-.268).

Also displayed in Table 6 are FMM results with alternative price measures. Instead of using the average price per ounce of ethanol, we enter price on each beer, wine and spirits in separate regressions. We find that the qualitative results remain the same. That is, the alcohol price is significant and negative only in Component 1. Also, for this group the elasticity is either greater than one (for beer and wine) or very close to one (for liquor). Specifically, the elasticities for Component 1 are -1.128, -1.515 and -.926 for beer, wine and liquor respectively.

Limitations. The results of this study are subject to several limitations. First, the number of drinks per day is the dependent variable. This measure does not capture all aspects of consumption (i.e. it does not capture days drinking). However, drinks per day may be a good proxy for the most problematic drinking as it is the number of drinks per occasion that engenders risky behavior. Some studies have also used drinks per day as their key measure of alcohol consumption (Kenkel 1996), while others have used this measure in a two-part model. Second, we are unable to differentiate between consumption of beer, wine, and spirits. Third, alcohol prices may be measured with some error and may be endogenous in the sense that higher demand may result in higher market prices[15] (Manning, et al. 1995). Fourth, for some of the behavioral economics variables, there were too few observations for our results to conclusively demonstrate the variables' effects. However, our preliminary results may prompt further interest in behavioral economics variables. In addition, the findings of this analysis are limited to older individuals while related research has typically focused on youths. Policymakers must consider the impact on the entire population.

Discussion

We find that there is important heterogeneity in the impact of alcohol price on alcohol consumption. For the majority of individuals, price is a significant determinant of demand for alcohol and these individuals are highly sensitive to price. For the other, smaller group, which consumes more alcohol, prices do not significantly affect consumption rates. Females, those with lower education, non-whites, and older individuals are more likely to be in the larger, high elasticity, low consumption group (Component 1). We also find that members of Component 1 are more likely to be risk averse, and have a shorter financial planning horizon.

Our results are in contrast to most of the earlier studies in this area, specifically because we identify a very price elastic group as well as find a group that does not respond significantly

[15] While endogeneity and measurement error may bias the results, this would apply to both components of the FMM.

to prices. As indicated above, Wagenaar et al. (2009) find the average alcohol price/tax elasticity to be -.03 across a large number of studies using individual level data[16]. Dave and Saffer (2007), find that chronic drinkers have a significant tax elasticity on the intensive margin of -.27. Manning et al. (1995) conclude that moderate drinkers are the most price sensitive, more so than heavier or lighter drinkers. Our estimate of a price elasticity of -1.6 is somewhat similar to the Manning et al. price estimate of -1.19 for the median drinker in quantile regressions; however, our estimate is based on Component 1 only.

Our finding that, among older adults, heavier drinkers are those who have inelastic demands for alcohol has important welfare implications for alcohol tax policy. The standard welfare approach would suggest (considering only efficiency conditions and ignoring distributional issues) that for most goods, higher taxes should be levied, ceteris paribus, on goods that have an inelastic demand. However, in the case of potentially harmful goods, sin taxes are levied in part to reduce the potential harms, external and internal, of consumption. For these drinkers, under several behavioral economic theories of addiction, taxes could increase welfare by serving as a precommitment device that serves to bolster weak self-control. However, our results suggest that the heavier drinkers are least likely to respond to the higher taxes, thus neither the externality nor 'internality'[17] justification for higher alcohol taxes is supported by our results. Note that those in Component 2 are heavier drinkers, but heavy drinking is only one factor in the complex set.

This paper adds to the literature in several dimensions. We apply the FMM estimation approach to estimating the price elasticity of demand. FMM allows us to unmask important, policy relevant heterogeneity in elasticity of demand for alcohol among older individuals. In

[16] The estimate is -.44 for studies using aggregate data.
[17] See for example Gruber, 2002, on the concept of internalities, intrapersonal impacts

addition, the FMM enable us to explore the impact of behavioral factors on the posterior odds of belonging to these heterogeneous latent classes. That is, we do not have to delineate only one, race, gender, or quantity of alcohol consumed, but rather can examine the multifaceted relationships underlying the latent classes. Figure 1 demonstrates graphically that the determination of the two components is not based on drinking status alone; many members of Component 2, for instance, consume 0 or 1 drink per day. By examining the posterior probability of being in each group, we can examine the impact of behavioral factors on classification into components.

While there is growing support for increased alcohol taxes at the political level, we are cautious about the welfare gains to increasing the tax on alcohol with regard to this older group of individuals. Prior to making a broad policy statement, we would want to apply our estimation approach to the full age spectrum. Nonetheless, we believe that this paper makes an important and provocative point and suggests a new approach to identifying heterogeneity in the price elasticity of demand for alcohol.

References

Abrams R, Alexopoulos G. (1977). Substance Abuse in the Elderly: Alcohol and Prescription Drugs. *Hospital and Community Psychiatry* 38(12):1285-1287.

Barsky RB, Juster FT, Kimball MS, Shapiro MD. (1997). Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study *Quarterly Journal of Economics* 112 (2): 537–579.

Bernheim, B. Douglas; Rangel, Antonio (2005). From Neuroscience to public policy: A New economic view of Addiction. Swedish Economic Policy Review of Addition.

Camerer, C., Issacharoff, S., Loewenstein, G., O'Donoghue, T., Rabin, M. (2003). Regulation for Conservatives: Behavioral Economics and the Case for "Aymmetric Paternalism". *University of Pennsylvania Law Review,* 151/3:1211-154.

Chaloupka FJ and Weschler H (1996). Binge drinking in college: the impact of price, availability, and alcohol control policies. *Contemporary Economic Policy* 14(4): 112–124.

Conway K, Deb P. (2005). Is Prenatal Care Really Ineffective? Or, is the 'Devil' in the Distribution? *Journal of Health Economics* 24, 489-513.

Dave D and Saffer H (2007) Risk Tolerance and Alcohol Demand among Adults and Older Adults. NBER Working Paper No. 13482

Deb P, Trivedi PK. (1997). Demand for Medical Care by the Elderly: A Finite Mixture Approach. *Journal of Applied Econometrics* 12, 313-336.

El-Gamal M, Grether D. (1995). Are People Bayesian? Uncovering Behaviorial Strategies. *Journal of the American Statistical Association* 90, 1137-1145.

French MT, Maclean JC (2006). Underage Alcohol Use, Delinquency, and Criminal Activity. *Health Economics* 15:1261-1281.

Gomberg E**.** (1982). Alcohol use and Problems among the Elderly. In: National Institute on Alcohol Abuse and Alcoholism, *Special Population Issues: Alcohol and Health Monograph 4*, DHHS Pub. No. (ADM) 82-1193. Washington, DC: Supt. of Docs., U.S. Govt. Print Off., pp. 263-290.

Grossman M, Chaloupka, FJ and Sirtalan, I. An empirical analysis of alcohol addiction: Results from the Monitoring the Future panels. *Economic Inquiry* 36(1):39–48, 1998.

Gruber J. (2002) Smoking's 'Internalities'. *Regulation* 25/4: 25-57.

Gruber J, Mullainathan S. (2002). Do Cigarette Taxes Make Smokers Happier? NBER Working Paper No. W887.

Gruber J, Koszegi B. (2000). Is Addiction 'Rational'? Theory and Evidence. *Quarterly Journal of Economics* 116(4): 1261-1303.

Hersch J. (2005) Smoking Restrictions as a Self-Control Mechanism. *Journal of Risk and Uncertainty* 31/1: 5-21

Heckman JJ, Singer B. (1984). A Method of Minimizing the Distributional Impact in Econometric Models for Duration Data. *Econometrica* 52:271-320.

Hurt R, Finlayson R, Morse R, Davis, L. (1988). Alcoholism in Elderly Persons: Medical Aspects and Prognosis of 216 Inpatients. *Mayor Clinic Proceedings* 64:753-760.

Juster FT, Suzman R. (1995). An Overview of the Health and Retirement Study. *Journal of Human Resources* 30:S7-56.

Kenkel DS. (1996). New Estimates of the Optimal Tax on Alcohol. *Economic Inquiry* 34: 296-319.

Kenkel DS. (2005) Are alcohol tax hikes fully passed through to prices: Evidence from Alaska. *AER,* 5(2):273-277.

Laird N. (1978). Non-parametric Maximum Likelihood Estimation of a Mixing Distribution. *Journal of the American Statistical Association* 73:805-811.

Leung SF, Phelps CE. (1993). My kingdom for a drink . . .? A review of estimates of the price sensitivity of demand for alcoholic beverages. In: Hilton, M.E. and Bloss, G., eds. *Economics and the Prevention of Alcohol-Related Problems*. NIAAA Research Monograph No. 25, NIH Pub. No. 93–3513. Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism, pp. 1–32.

Lindsay BJ. (1995). Mixture Models: Theory, Geometry, and Applications, NSF-CBMS Regional Conference Series in Probability and Statistics, Vol. 5, IMS-ASA.

McLachlan GJ, Peel D. (2000) *Finite Mixture Models.* New York: John Wiley.

Manning WG, Blumberg L and Moulton, LH. (1995) The demand for alcohol: The differential response to price. *Journal of Health Economics* 14(2):123–148.

Manning, WG, Keeler EB, Newhouse JP, Sloss EM, Wasserman J. (1989). The Taxes of Sin: Do Smokers & Drinkers Pay Their Way. *Journal of the American Medical Association* 261(11):1604-1609.

Morduch J, Stern HS. (1997) Using Mixture Models to Detect Sex Bias in Health Outcomes in Bangladesh. *Journal of Econometrics* 77, 259-276.

O'Donoghue, T. Rabin, M. (2003). Studying Optimal Paternalism, Illustrated with a Model of Sin Taxes. AER Papers and Proceedings, 93/2: 191.

Ornstein, S.I. (1980) Control of alcohol consumption through price increases. *Journal of Studies on Alcohol* 41: 807–818

Ornstein, S.I. & Levy, D. (1983) Price and income elasticities and the demand for alcoholic beverages. In Galanter, M. (ed.): Recent developments in alcoholism, pp. 303–345. New York: Plenum.

Radloff LS. (1977) The CES-D Scale: A Self-report Depression Scale for Research in the General Population, *Applied Psychological Measurement* 1(3): 385–401.

Saffer H, Dave D. (2005) Mental Illness and the Demand for Alcohol, Cocaine, and Cigarettes, *Economic Inquiry*, Oxford University Press 43(2): 229-246.

Sloan FA, Reilly BA, Schenzler C. (1994). Effects of Prices, Civil and Criminal Sanctions, and Law Enforcement on Alcohol-related Mortality. *Journal of Studies on Alcohol* 55:454–465.

Vestal R, McGuire E, Tobin J, Andres R, Norris M, Mesey, E. Aging and Alcohol Metabolism. *Clinical Pharmacology and Therapeutics* 21(3):343-354, 1977.

Vogel-Sprott **V,** Barrett P. (1984). Age, Drinking Habits and the Effects of Alcohol. *Journal of Studies on Alcohol* 48(6):517-521.

Wagenaar AC, Salois MJ, Komro KA. (2009). Effects of Beverage Alcohol Price and Tax Levels on Drinking: A Meta-Analysis of 1003 Estimates from 112 Studies. *Addiction* 104: 179-190.

Wang P, Cockburn IM, Puterman ML. (1998) Analysis of Patent Data - A Mixed Poisson Regression Model Approach. *Journal of Business and Economic Statistics* 16, 27-36.

Wedel M, Desarbo WS, Bult JR, Ramaswamy V. (1993). A Latent Class Poisson Regression Model for Heterogeneous Count Data. *Journal of Applied Econometrics 8, 397-411.*

Williams, M**.** (1984). Alcohol and the Elderly: An Overview. *Alcohol Health & Research World* 8(3): 3-9, 52.

Young, DJ and Bielinska-Kwapsisz A (2002) Alcohol Taxes and beverage prices. National Tax Journal. LV(1): 57-74.

Table 1: Summary Statistics

	Mean	Std. Dev.
Drinks per day	0.616	1.346
Average price per ounce of ethanol	8.308	1.063
Age 0 – 45	0.014	
Age 46 – 55	0.147	
Age 56 – 65	0.377	
Age 66 – 75	0.256	
Age 76 – 85	0.155	
Age 86+	0.051	
Black	0.151	
Other race	0.031	
Male	0.419	
Height	1.688	0.101
Married or partnered	0.675	
Household income (in 10,000 1992 USD)	4.463	7.143
Years of education	12.073	3.253
Binge drinking	1.325	8.242
CESD score	1.512	1.941
Risk aversion (imputed)	3.301	0.989
Financial planning horizon (imputed)	2.970	1.131
Observations	71,802	

Note that the number of observations is not the same for all variables

Figure 1: Predicted Densities of the 2 Components

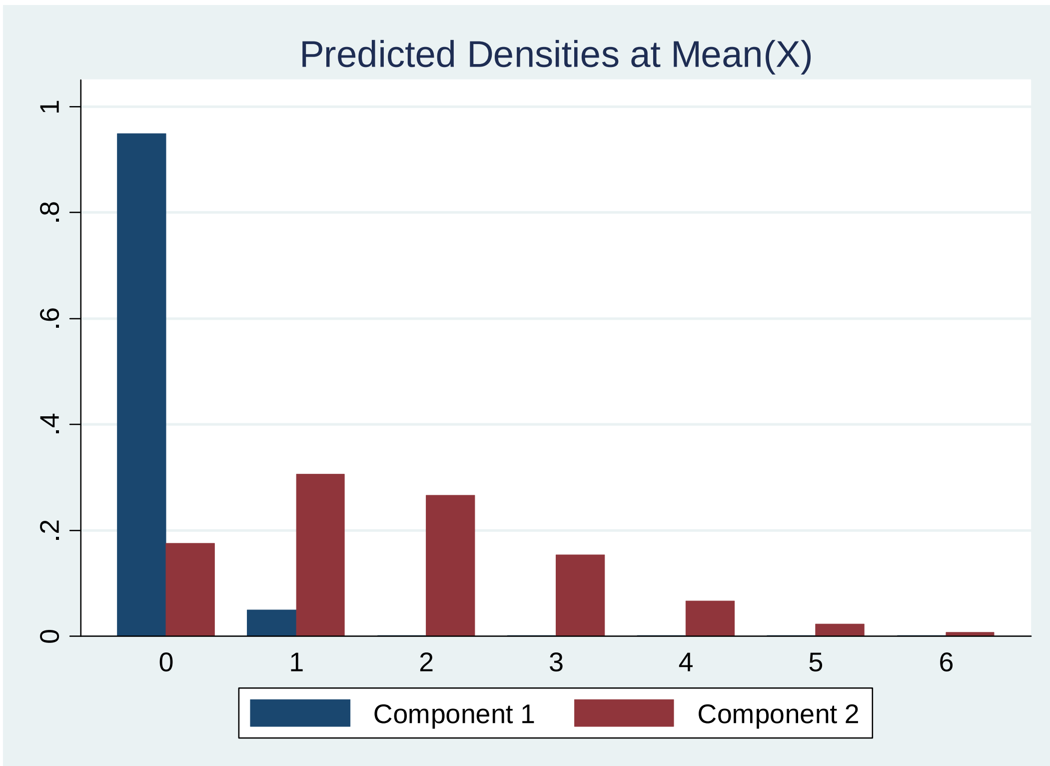

Table 2. Price Elasticity of the Number of Drinks

	Poisson	FMM Component1	FMM Component2	FMM Prior Probability of Component 1
Log average price	-0.286***	-1.600***	0.0351	
	(0.0918)	(0.345)	(0.134)	
Age 46-55	-0.559***	0.0696	-0.405***	0.274*
	(0.0632)	(0.367)	(0.112)	(0.162)
Age 56-65	-0.771***	0.183	-0.532***	0.487***
	(0.0619)	(0.368)	(0.112)	(0.161)
Age 66-75	-1.061***	0.356	-0.721***	0.757***
	(0.0628)	(0.369)	(0.112)	(0.164)
Age 76-85	-1.446***	0.158	-1.084***	0.846***
	(0.0648)	(0.368)	(0.117)	(0.175)
Age 86+	-2.040***	-0.549	-1.573***	0.943***
	(0.0824)	(0.426)	(0.187)	(0.277)
Height	0.231**	1.021*	0.0533	-0.140
	(0.117)	(0.538)	(0.213)	(0.363)
Years of education	0.0334***	0.248***	-0.0580***	-0.115***
	(0.00281)	(0.0200)	(0.00490)	(0.0111)
Black	-0.332***	-2.636***	-0.122***	0.150**
	(0.0280)	(0.680)	(0.0443)	(0.0754)
Other race	-0.180***	-1.980***	0.0913	0.346**
	(0.0665)	(0.760)	(0.124)	(0.149)
Male	0.898***	-0.0980	0.460***	-0.897***
	(0.0229)	(0.114)	(0.0426)	(0.0756)
Constant	-0.189	-3.171***	1.447***	2.488***
	(0.281)	(1.172)	(0.466)	(0.599)
Observations	71,802	71,802	71,802	71,802
Mean Predicted y		0.129	1.879	
% in Component 1		75.4	24.6	

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Regressions also include census region of residence and year dummies

Table 3. Determinants of the Posterior Probability of Component 1

	(1)	(2)	(3)	(4)	(5)
Age 46-55	0.0502***	0.0509***	0.0411	0.0498***	0.0478***
	(0.0133)	(0.0134)	(0.0390)	(0.0142)	(0.0153)
Age 56-65	0.0918***	0.0910***	0.104***	0.0883***	0.110***
	(0.0131)	(0.0131)	(0.0380)	(0.0140)	(0.0150)
Age 66-75	0.141***	0.139***	0.142***	0.139***	0.166***
	(0.0131)	(0.0131)	(0.0383)	(0.0143)	(0.0150)
Age 76-85	0.153***	0.151***	0.134***	0.146***	0.175***
	(0.0131)	(0.0133)	(0.0390)	(0.0155)	(0.0151)
Age 86+	0.159***	0.157***	0.188***	0.152***	0.182***
	(0.0134)	(0.0136)	(0.0407)	(0.0287)	(0.0156)
Height	-0.0383**	-0.0305*	-0.120**	-0.0102	-0.0341
	(0.0179)	(0.0179)	(0.0581)	(0.0289)	(0.0210)
Years of education	-0.0191***	-0.0177***	-0.0185***	-0.0174***	-0.0210***
	(0.000374)	(0.000392)	(0.00125)	(0.000678)	(0.000455)
Black	0.0258***	0.0230***	-0.0225	0.0116**	0.0186***
	(0.00366)	(0.00371)	(0.0150)	(0.00588)	(0.00435)
Other race	0.0494***	0.0454***	0.0571**	0.0328***	0.0509***
	(0.00739)	(0.00744)	(0.0261)	(0.0112)	(0.00869)
Male	-0.169***	-0.169***	-0.146***	-0.180***	-0.172***
	(0.00366)	(0.00373)	(0.0112)	(0.00601)	(0.00431)
Married		0.00865***			
		(0.00290)			
Log income		-0.0108***			
		(0.000970)			
Locus of control			-0.0247***		
			(0.00328)		
Forward looking			-0.00646*		
			(0.00354)		
Risk aversion				0.0102***	
				(0.00198)	
Financialplanning horizon				-0.0126***	
				(0.00188)	
Binge drinking					-0.0113***
					(0.000215)
CESD score					0.00865***
					(0.000724)
Constant	0.970***	1.046***	1.133***	0.903***	0.934***
	(0.0325)	(0.0334)	(0.101)	(0.0508)	(0.0380)
Observations	71,802	71,713	8,542	32,989	53,227
R-squared	0.121	0.123	0.108	0.104	0.214

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: FMM Models

	Component1	Component2	Prior Probability of Component 1	Component1	Component2	Prior Probability of Component 1
Log average price	-1.464***	0.131		-1.522***	0.0388	
	(0.442)	(0.187)		(0.344)	(0.136)	
Risk aversion	-0.113***	0.00190				
	(0.0426)	(0.0153)				
Financial planning horizon	0.248***	-0.0614***				
	(0.0436)	(0.0196)				
Blood alcohol count law				0.0715	0.0123	
				(0.0783)	(0.0276)	
License revocation law				0.0558	-0.0452	
				(0.125)	(0.0448)	
Zero tolerance law				-0.224*	-0.0698*	
				(0.132)	(0.0395)	
Age 46-55	0.159	-0.416***	0.261	0.0725	-0.406***	0.274*
	(0.380)	(0.123)	(0.165)	(0.361)	(0.112)	(0.161)
Age 56-65	0.338	-0.539***	0.482***	0.186	-0.533***	0.487***
	(0.385)	(0.125)	(0.167)	(0.362)	(0.111)	(0.160)
Age 66-75	0.507	-0.752***	0.795***	0.363	-0.721***	0.761***
	(0.403)	(0.130)	(0.190)	(0.363)	(0.112)	(0.163)
Age 76-85	0.491	-1.122***	0.760***	0.167	-1.084***	0.851***
	(0.422)	(0.150)	(0.255)	(0.363)	(0.117)	(0.175)
Age 86+	0.610	0.732***	3.845***	-0.547	-1.575***	0.941***
	(0.559)	(0.139)	(0.961)	(0.424)	(0.186)	(0.277)
Height	0.421	0.106	0.0458	1.031*	0.0520	-0.137
	(0.730)	(0.266)	(0.499)	(0.537)	(0.213)	(0.363)
Years of education	0.255***	-0.0591***	-0.0815***	0.247***	-0.0579***	-0.114***
	(0.0263)	(0.00673)	(0.0149)	(0.0200)	(0.00491)	(0.0111)
Black	-3.162*	-0.156***	-0.0398	-2.653***	-0.121***	0.148**
	(1.684)	(0.0576)	(0.103)	(0.691)	(0.0445)	(0.0756)
Other race	-1.929*	0.0707	0.141	-1.980***	0.0922	0.346**
	(1.068)	(0.180)	(0.206)	(0.753)	(0.123)	(0.149)
Male	-0.0765	0.455***	-0.920***	-0.104	0.459***	-0.901***
	(0.154)	(0.0536)	(0.105)	(0.115)	(0.0426)	(0.0758)
Constant	-3.098*	1.449**	1.835**	-3.300***	1.508***	2.484***
	(1.656)	(0.661)	(0.818)	(1.175)	(0.472)	(0.599)
Observations	32,989	32,989	32,989	71,802	71,802	71,802
Mean predicted y	0.159	2.084		0.129	1.879	
% in component 1	0.723	0.277		0.754	0.246	

*** p<0.01, ** p<0.05, * p<0.1

Robust standard errors in parentheses

Regressions also include census region of residence and year dummies

Table 5: Selection Criteria for Various Models

	Log-likelihood	d.o.f	AIC	BIC
Poisson	-82363.43	24	164774.9	164995.2
2 component, constant probability	-70321.87	49	140741.7	141191.6
3 component, constant probability	Not converging			
2 component, variable probability	-69550.2	59	139218.4	139760.1
Risk/Planning Model	-35478.15	63	71082.29	71611.74
Laws Model	-69543.24	65	139216.5	139813.3
Only drinkers sample (2 component, variable probability)	-58492.72	59	117103.4	117615

Table 6: Comparison of Price Elasticity Estimates across Models

	Overall	Component 1	Component 2	Component 3	N
Poisson	-0.286*** (0.0918)				71,802
2 component, constant probability		-0.997*** (0.245) [0.754]	-0.023 (0.142) [0.246]		71,802
3 component, constant probability	Not converging				
2 component, variable probability		-1.600*** (0.345) [0.754]	0.0351 (0.134) [0.246]		71,802
Beer		-1.128*** (0.285) [0.754]	0.0282 (0.107) [0.246]		71,802
Wine		-1.515*** (0.270) [0.755]	0.0816 (0.117) [0.245]		71,802
Liquor		-0.926*** (0.228) [0.754]	0.00649 (0.133) [0.246]		71,802
Only drinkers sample (2 component, variable probability)		-0.630*** (0.224) [0.604]	0.152 (0.149) [0.396]		43,098

Robust standard errors in parentheses, proportion in component in square brackets
*** p<0.01, ** p<0.05, * p<0.1