

**A SUMMARY OF SOME ESTIMATORS OF DYNAMIC
PANEL DATA MODELS AND THEIR APPLICATIONS**

by

Zhen Ma

A dissertation submitted to the Graduate Faculty in Economics in partial fulfillment of
the requirements for the degree of Doctor of Philosophy,
The City University of New York

2012

This manuscript has been read and accepted for the Graduate Faculty in Economics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Professor Michael Grossman

Date

Chair of Examining Committee

Professor Merih Uctum

Date

Executive Officer

Professor Michael Grossman

Professor Theodore Joyce

Professor Wim Vijverberg

Supervisory Committee

CITY UNIVERSITY OF NEW YORK

Abstract

A SUMMARY OF SOME ESTIMATORS OF DYNAMIC PANEL DATA MODELS AND THEIR APPLICATIONS

by

Zhen Ma

Advisor: Professor Michael Grossman

This thesis consists of two chapters. Chapter one summarizes three estimators of dynamic panel data models: Generalized Method of Moments (GMM) with fixed effects, Wooldridge Conditional Maximum Likelihood (CML) with random effects and a Maximum Simulated Likelihood (MSL) random effects dynamic probit. Chapter two presents their applications and empirical findings. I examine the impact of the large price increases in cigarettes after the Master Settlement Agreement (MSA) on drinking behavior using data from the Panel Study of Income Dynamics (PSID). Alcohol consumption, drinking participation and heavy drinking participation (three or more drinks per day) are considered for the full sample, as well as for sub-samples stratified by age group and gender. Estimation results are relatively stable across estimators. I find that the cross-price effects of cigarettes on alcohol consumption are insignificant showing that averaging on all consumption levels, the number of drinks consumed per day is not affected by the increases in cigarette prices; and that the cross-price effects of cigarettes on drinking participation are mostly positive and significant, indicating drinking is an economic substitute for smoking; also, cigarette prices do not affect heavy drinking prevalence.

Acknowledgments

I would like to express my sincere gratitude to Professor Michael Grossman, Professor Ted Joyce and Professor Wim Vijverberg for their most valuable guidance and support during the six years of my studies in the program and my job search. I would also like to thank and express my gratitude to Professor Elizabeth Roistacher who has been my mentor for the past ten years.

Table of Contents

Chapter One.....	1
1.1 Prototypical Dynamic Panel Data Model.....	1
1.2 An Overview of GMM.....	3
1.2.1 Difference GMM.....	3
1.2.2 System GMM.....	6
1.3 Wooldridge Conditional Maximum Likelihood (CML).....	9
1.4 Maximum Simulated Likelihood (MSL).....	11
Chapter Two.....	14
2.1 Introduction.....	14
2.2 Background and Literature.....	18
2.3 Conceptual Framework and Empirical Model.....	20
2.4 Data.....	22
2.5	26
2.5.1 Estimation of Equation (2.1).....	26
2.5.2 Estimation Results.....	29
2.5.3 Alcohol Consumption Using GMM.....	30
2.5.4 Drinking Participation.....	31
2.5.5 Heavy Drinking Participation.....	33
2.6 Discussion.....	34
Tables.....	37
References.....	48

List of Tables

Table 1. Alcohol and Cigarette Use Descriptive Statistics.....	37
Table 2. Independent Variables Descriptive Statistics.....	38
Table 3. OLS and FE Estimates for Alcohol Consumption.....	39
Table 4. GMM Estimates for Alcohol Consumption with Lagged Smoking Participation....	40
Table 5. GMM Estimates for Alcohol Consumption with Lagged Cigarette Consumption....	41
Table 6. OLS, FE and RE Estimates for Drinking Participation.....	42
Table 7. CML Estimates for Drinking Participation with Lagged Smoking Participation.....	43
Table 8. CML Estimates for Drinking Participation with Lagged Cigarette Consumption....	44
Table 9. MSL Estimates for Drinking Participation.....	45
Table 10. CML Estimates for Heavy Drinking Participation.....	46
Table 11. MSL Estimates for Heavy Drinking Participation.....	47

Chapter One

Panel (longitudinal) data offer three major advantages as summarized by Cameron and Trivedi (2005). The first advantage is increased precision in estimation. By pooling several time periods of data for each individual, panel data give the researcher more observations, increasing the degrees of freedom and reducing the collinearity among explanatory variables (Hsiao, 2003). A second advantage is the possibility of mitigating the omitted variables bias problem when it is difficult to find valid instrumental variables using cross sectional data. Omitted variable bias arises when the unobserved individual heterogeneity is correlated with regressors. A third attraction of panel data is the capability of estimating dynamic models. Dynamic models are of interest in a wide range of economic applications. Even when coefficients on lagged dependent variables are not of direct interest, allowing dynamics in the underlying process may be crucial for recovering consistent estimates for other parameters (Bond 2002). However, difficulty arises in estimating dynamic models due to correlation between time invariant unobserved individual heterogeneity and the lagged dependent variable. This chapter summarizes three consistent estimators of dynamic panel data models, all of which can be easily implemented in Stata: Generalized Method of Moments (GMM) with fixed effects, Wooldridge Conditional Maximum Likelihood (CML) with random effects and a Maximum Simulated Likelihood (MSL) random effects dynamic probit. Chapter two presents their applications and empirical findings.

1.1 Prototypical Dynamic Panel Data Model

Consider a general first-order autoregressive panel data model specified as

$$(1.1) \quad y_{it} = \alpha y_{i,t-1} + \beta x_{i,t} + \varepsilon_{i,t}, \quad i = 1, 2, \dots, n; \quad t = 2, 3, \dots, T, \quad \text{where } \varepsilon_{i,t} = u_i + v_{i,t}.$$

The error term $\varepsilon_{i,t}$ consists of the unobserved individual fixed effects u_i , and the idiosyncratic disturbances $v_{i,t}$. u_i and $v_{i,t}$ are assumed to have zero unconditional means and not correlated with each other. In addition, the idiosyncratic disturbances $v_{i,t}$ are not correlated across individuals. That is, $E(u_i) = 0$, $E(v_{i,t}) = 0$, $E(u_i v_{i,t}) = 0$, for $i = 1, 2, \dots, n$ and $t = 2, 3, \dots, T$. The lagged dependent variable $y_{i,t-1}$ is a predetermined variable: it is not influenced by the current disturbances $v_{i,t}$, but is affected by the past ones. The regressor $x_{i,t}$ is a vector of independent variables that can include strictly exogenous, endogenous or predetermined variables.

Estimating equation (1.1) with OLS will result in biased and inconsistent estimation because the factors that determine past period of y_i persist in this period, and are captured by the individual fixed effect term u_i . One would attempt to employ the fixed effects estimation to purge the fixed effects. It first performs a within groups or time demean transformation to each variable and then use OLS on the transformed data. However, fixed effects estimator does not eliminate biases from dynamic panel data when the time dimension T is small (Bond 2002). The

transformed lagged dependent variable is $y_{i,t-1}^* = y_{i,t-1} - \frac{1}{T-1}(y_{i,1} + \dots + y_{i,T-1})$ and the

transformed error term is $v_{i,t}^* = v_{i,t} - \frac{1}{T-1}(v_{i,2} + \dots + v_{i,t-1} + \dots + v_{i,T})$. $\frac{-y_{i,t}}{T-1}$ in $y_{i,t-1}^*$ is

negatively correlated with $v_{i,t}$ in $v_{i,t}^*$ and $y_{i,t-1}$ in $y_{i,t-1}^*$ is negatively correlated with $\frac{-v_{i,t-1}}{T-1}$ in $v_{i,t}^*$

. Nickell (1981) points out that these negative correlations dominate other positive correlations between $y_{i,t-1}^*$ and $v_{i,t}^*$ such that the correlation between the transformed lagged dependent variable and the transformed error term remains negative. Accordingly, fixed effects estimates will be downward biased and inconsistent when the number of time periods T is small.

1.2 An Overview of GMM

Among all estimators that solve the dynamic panel data bias problem, difference generalized method of moments (GMM) estimator and system generalized method of moments (GMM) estimator are most popular ones.

1.2.1 Difference GMM

Difference GMM was first proposed by Holtz-Eakin, Newey and Rosen (1988) and enriched by Arellano and Bond (1991). It begins by transforming equation

$$(1.2) \quad y_{it} = \alpha y_{i,t-1} + \beta x_{i,t} + \varepsilon_{i,t}, \quad i = 1, 2, \dots, n; \quad t = 2, 3, \dots, T, \quad \text{where } \varepsilon_{i,t} = u_i + v_{i,t}$$

into first differences to eliminate the fixed effects:

$$(1.3) \quad \Delta y_{i,t} = \alpha \Delta y_{i,t-1} + \beta \Delta x_{i,t} + \Delta v_{i,t}$$

Note that the regressors in first differences are still correlated with the error term since $y_{i,t-1}$ is correlated with $v_{i,t-1}$, $\Delta y_{i,t-1}$ is correlated with $\Delta v_{i,t}$ ($\Delta y_{i,t-1} = y_{i,t-1} - y_{i,t-2}$ and $\Delta v_{i,t} = v_{i,t} - v_{i,t-1}$); with the same line of reasoning, first differences of any endogenous or predetermined variables in $x_{i,t}$ are now correlated with the error term. In addition to first differencing, use the lagged levels of right-hand side predetermined or endogenous variables from equation (1.2) as instruments for their first-differenced form in equation (1.3). Suppose in equation (1.2), $y_{i,t-1}$ is predetermined and $x_{i,t}$ is a strictly exogenous variable. $y_{i,t-2}$ can be a valid instrument for $\Delta y_{i,t-1}$ since it is correlated with $\Delta y_{i,t-1} = y_{i,t-1} - y_{i,t-2}$, but orthogonal to $\Delta v_{i,t} = v_{i,t} - v_{i,t-1}$. All previous realizations of y_i can be potential instruments. $\Delta x_{i,t}$ instrument themselves¹. In general, there

¹ $\Delta x_{i,t}$ can also be instrumented by $x_{i,t}$.

are $m + k(T - 2)$ linear orthogonality conditions where $m = (T - 1)(T - 2)/2$ and k is the number of exogenous variables (in this case 1). That is

$$(1.4) \quad E(y_{i,t-s} \Delta v_{i,t}) = 0 \quad t = 3, 4, \dots, T \text{ and } s = 2, 3, \dots, t - 1 \text{ and}$$

$$(1.5) \quad E(\Delta x_{i,t} \Delta v_{i,t}) = 0 \quad t = 3, 4, \dots, T$$

These conditions hold only under the absence of serial correlation in the idiosyncratic disturbances $v_{i,t}$. In addition, one more assumption regarding the initial conditions $y_{i,1}$ is needed

$$(1.6) \quad E(y_{i,1} v_{i,t}) = 0 \quad i = 1, 2, \dots, N \text{ and } t = 2, 3, \dots, T$$

The moment conditions (1.4) for each individual can be written more compactly as:

$$(1.7) \quad E(Z'_{di} \Delta U_i) = 0$$

where Z_{di} is the $(T - 2) \times (m + (T - 2))$ block diagonal matrix (with s blocks) given by:

$$(1.8) \quad Z_{di} = \begin{pmatrix} y_{i1} & 0 & \dots & 0 & 0 & \Delta x_{i3} \\ 0 & y_{i2} y_{i3} & \dots & & 0 & \Delta x_{i4} \\ \cdot & \cdot & \dots & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & y_{i2} \dots y_{iT-3} & 0 & \Delta x_{iT-1} \\ 0 & 0 & 0 & 0 & y_{i1} \dots y_{iT-2} & \Delta x_{iT} \end{pmatrix}$$

and ΔU_i is the $(T - 2)$ vector:

$$(1.9) \quad \Delta U_i = \begin{bmatrix} \Delta v_{i3} \\ \Delta v_{i4} \\ \vdots \\ \Delta v_{iT} \end{bmatrix}$$

Stacking matrix Z'_{di} and vector $\Delta U'_i$ across individuals, we get the $(m + (T - 2)) \times N(T - 2)$

matrix $Z'_d = (Z'_{d1}, Z'_{d2}, \dots, Z'_{dN})$ and $N(T - 2)$ vector $\Delta U = (\Delta U'_1, \Delta U'_2, \dots, \Delta U'_N)'$, respectively.

Based on the moment conditions (1.7), GMM estimation chooses $\hat{\alpha}$ and $\hat{\beta}$ to minimize the quadratic distance $\Delta U'Z_d W_N Z_d' \Delta U$, W_N being a symmetric positive definite weighting matrix²

$$(1.10) \quad \hat{B}^d_{1-step} = (\Delta \tilde{x}' Z_d W_N Z_d' \Delta \tilde{x})^{-1} \Delta \tilde{x}' Z_d W_N Z_d' \Delta y$$

where $\hat{B} = (\hat{\alpha}, \hat{\beta})'$, $\tilde{x} = (y_{-1}, x)$, $\Delta y'$ is the $N(T-2)$ vector $(\Delta y_3, \Delta y_4, \dots, \Delta y_T)$ and $\Delta \tilde{x}'$ is the

$$2 \times N(T-2) \text{ matrix } \begin{pmatrix} \Delta y_2 & \Delta y_3 & \dots & \Delta y_{T-1} \\ \Delta x_3 & \Delta x_4 & \dots & \Delta x_T \end{pmatrix}. \text{ The choice of the weighting matrix is}$$

independent of the consistency of the GMM estimator for large N. Equation (1.10) is called 1-step difference GMM estimator when an arbitrary weighting matrix is used.

To obtain the efficient GMM estimator, the optimal weighting matrix should be used

$$(1.11) \quad \hat{S}_N = (N^{-1} \sum_{i=1}^N Z'_{di} \Delta \hat{U}_i \Delta \hat{U}'_i Z_{di})^{-1}$$

where \hat{U}_i are the residuals from estimation of \hat{B}^d_{1-step} . Hence the 2-step difference GMM estimator

$$(1.12) \quad \hat{B}^d_{2-step} = (\Delta \tilde{x}' Z_d \hat{S}_N Z_d' \Delta \tilde{x})^{-1} \Delta \tilde{x}' Z_d \hat{S}_N Z_d' \Delta y$$

However, different choices of the weighting matrix W_N will yield different point estimates. Arellano and Bond (1991) suggests using

$$(1.13) \quad W_N = (N^{-1} \sum_{i=1}^N Z'_{di} H Z_{di})^{-1}$$

to yield the initial consistent estimator, where H is the $(T-2)$ square matrix given by

² The superscript d indicates difference GMM.

$$(1.14) \quad H = \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ 0 & -1 & 2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \cdot \\ 0 & 0 & 0 & \dots & 2 \end{pmatrix}$$

Then, the 2-step difference GMM estimator is obtained by using the residuals from the first step as $\Delta \hat{U}_i$ in (1.11).

1.2.2 System GMM

Blundell and Bond (1998) have shown that difference GMM estimator can have the weak instrument problem under two circumstances: 1) if y_i is close to a random walk process, or 2) as $\text{var}(u_{i,t})/\text{var}(v_{i,t})$, the variance of u_i relative to that of v_i , increases. Weak instruments inflate the bias of 2SLS estimator, even asymptotically, in the direction of the OLS estimator. In addition, weak instruments make 2SLS estimators less efficient than OLS estimators.³

Arellano and Bover (1995) and Blundell and Bond (1998) augment difference GMM by utilizing moment conditions for the level equation under an additional assumption on the initial condition

$$(1.15) \quad E(\Delta y_{i,2} u_i) = 0$$

Besides first differencing the regressors to eliminate fixed effects, they suggest first differencing the past levels (now exogenous to the fixed effects) to instrument the regressors in levels.

Blundell and Bond (1998) show that lagged difference as instruments remain informative even when α approaches unity or when $\text{var}(u_{i,t})/\text{var}(v_{i,t})$ becomes large. Returning to equation

(1.1), $\Delta y_{i,t-1} = y_{i,t-1} - y_{i,t-2}$ can be a valid instrument for $y_{i,t-1}$ since it is correlated with $y_{i,t-1}$, but

³ 2SLS estimator is GMM with an efficient weighting matrix under homoskedasticity.

orthogonal to u_i and $v_{i,t}$. Analogous to the case of difference GMM, there are $m + k(T - 2)$ linear orthogonality conditions where $m = (T - 1)(T - 2)/2$ and k is the number of exogenous variables (in this case 1). That is

$$(1.16) \quad E(\Delta y_{i,t-s} v_{i,t}) = 0 \quad t = 3, 4, \dots, T \text{ and } s = 2, 3, \dots, t-1 \text{ and}$$

$$(1.17) \quad E(x_{i,t} v_{i,t}) = 0 \quad t = 3, 4, \dots, T$$

This is valid under the assumption that changes in instruments are uncorrelated with the fixed effects u_i . Again, their validity also depends on the assumption of no serial correlation in $v_{i,t}$.

In matrix notation:

$$(1.18) \quad E(Z'_i U_i) = 0 \quad \text{where } Z_i \text{ is the } (T - 2) \times (m + (T - 2)) \text{ matrix given by:}$$

$$(1.19) \quad Z_i = \begin{pmatrix} \Delta y_{i1} & 0 & \dots & 0 & x_{i3} \\ 0 & \Delta y_{i1} \Delta y_{i2} & \dots & 0 & x_{i4} \\ \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & 0 & 0 & \Delta y_{i1} \dots \Delta y_{iT-3} & x_{iT-1} \\ 0 & 0 & 0 & 0 & \Delta y_{i1} \dots \Delta y_{iT-2} x_{iT} \end{pmatrix}$$

and U_i is the $(T - 2)$ vector:

$$(1.20) \quad U_i = \begin{bmatrix} v_{i3} \\ v_{i4} \\ \vdots \\ v_{iT} \end{bmatrix}$$

Let $Z'_l = (Z'_{i1}, Z'_{i2}, \dots, Z'_{iN})$ and $U = (U'_1, U'_2, \dots, U'_N)'$. The GMM estimator based on these moment conditions, called level GMM, is given by⁴

⁴ The superscript l indicates level GMM.

$$(1.21) \quad \hat{B}' = (\tilde{x}'Z_l'W_NZ_l'\tilde{x})^{-1}\tilde{x}'Z_l'W_NZ_l'y \quad \text{where } y' \text{ is the } N(T-2) \text{ vector } (y_3, y_4, \dots, y_T) \text{ and } \tilde{x}'$$

is the $2 \times N(T-2)$ matrix $\begin{pmatrix} y_2 & y_3 & \dots & y_{T-1} \\ x_3 & x_4 & \dots & x_T \end{pmatrix}$.

Combining the two sets of moment conditions, we get

$$(1.22) \quad E(Z_s'Q) = 0 \quad \text{where } Z_s \text{ is the } 2N(T-2) \times 2(m+(T-2)) \text{ matrix}$$

$$(1.23) \quad Z_s = \begin{pmatrix} Z_d & 0 \\ 0 & Z_l \end{pmatrix} \quad \text{and } Q \text{ is the } 2N(T-2) \text{ vector}$$

$$(1.24) \quad Q = \begin{pmatrix} \Delta U \\ U \end{pmatrix}$$

The system of difference and level equations is estimated as one equation based on these moment conditions. The resulting estimator is called system GMM estimator, given by⁵

$$(1.25) \quad \hat{B}^s = (\tilde{X}'Z_s'W_NZ_s'\tilde{X})^{-1}\tilde{X}'Z_s'W_NZ_s'Y$$

in which Y is the $2N(T-2)$ stacked vector:

$$(1.26) \quad Y = \begin{pmatrix} \Delta y \\ y \end{pmatrix} \quad \text{where } \Delta y' \text{ is the } N(T-2) \text{ vector } (\Delta y_3, \Delta y_4, \dots, \Delta y_T) \text{ and } y' \text{ is the } N(T-2)$$

vector (y_3, y_4, \dots, y_T) , and \tilde{X} is the $2N(T-2) \times 2$ stacked matrix

$$(1.27) \quad \tilde{X} = \begin{pmatrix} \Delta \tilde{x} \\ \tilde{x} \end{pmatrix}$$

where $\Delta \tilde{x}$ is the $N(T-2) \times 2$ matrix $\begin{pmatrix} \Delta y_2 & \Delta y_3 & \dots & \Delta y_{T-1} \\ \Delta x_3 & \Delta x_4 & \dots & \Delta x_T \end{pmatrix}'$ and \tilde{x} is the $N(T-2) \times 2$ matrix

$\begin{pmatrix} y_2 & y_3 & \dots & y_{T-1} \\ x_3 & x_4 & \dots & x_T \end{pmatrix}'$. W_N is a $2N(T-2) \times 2N(T-2)$ stacked weighting matrix. 1-step and 2-

⁵ The superscript s indicates system GMM.

step system GMM estimators are obtained in the similar fashion as in the case of difference GMM discussed previously. Blundell and Bond (1998) demonstrate in simulations that system GMM achieves dramatic efficiency gains over difference GMM especially at high values of α .

In sum, Arellano and Bond (1991) difference GMM and Arellano and Bover (1995) / Blundell and Bond (1998) system GMM estimators are both designed for panel data of few time periods T and large number of cross sectional groups N . Usually individual fixed effects are present and idiosyncratic disturbances are not serially correlated. Some regressors are not strictly exogenous and instruments are drawn from within the data set.⁶

1.3 Wooldridge Conditional Maximum Likelihood (CML)

A random effects model assumes that the unobservable individual effects are random variables that are distributed independently of the regressors.

In a discrete choice context, Heckman (1981) and Wooldridge (2005) are the two standard estimators for the dynamic random effects model. Heckman (1981) suggests specifying the joint probability distribution of the response variable including the initial observation $y_{i,1}$ conditional on the observed exogenous variables and unobserved heterogeneity by approximating the conditional distribution of $y_{i,1}$. The joint probability of all observed y can be written as

$$(1.28) \quad f(y_{i,1}, y_{i,2}, \dots, y_{i,T} | x_{i,t}, u_i) = f(y_{i,2}, y_{i,3}, \dots, y_{i,T} | y_{i,1}, x_{i,t}, u_i) g(y_{i,1} | x_{i,t}, u_i).$$

⁶ Exogenous variables are instrumented by themselves. IV style instruments from outside the model are also allowed to instrument exogenous variables.

If an assumption of the conditional density of the first observation $g(y_{i,1} | x_{i,t}, u_i)$ is made, and if the density for u_i given $x_{i,t}$ is specified as normal, then the unobserved u_i can be integrated out and the joint probability of all observed y can be approximated,

$$(1.29) \quad f(y_{i,1}, y_{i,2}, \dots, y_{i,T} | x_{i,t}, u_i) = \int f(y_{i,2}, y_{i,3}, \dots, y_{i,T} | y_{i,1}, x_{i,t}, u_i) g(y_{i,1} | x_{i,t}, u_i) h(u_i | x_i) du_i$$

Heckman suggested a probit model for the initial period $y_{i,1} = 1(\pi z_{i,1} + \kappa_i > 0)$, where

$\kappa_i = \theta u_i + v_{i,1}$, $z_{i,1}$ includes $x_{i,1}$ and additional pre-sample variables. The drawback of the

Heckman estimator is that there is no standard software available to implement it.

Wooldridge (2005) considers the distribution $f(y_{i,2}, y_{i,3}, \dots, y_{i,T})$ conditional on $y_{i,1}$ and proposes a much simpler approach that can be implemented with standard random effects probit software. Instead of approximating the distribution of $y_{i,1}$ conditional on x_i and u_i to specify the joint probability of y_i of all periods, the distribution of u_i conditional on $y_{i,1}$ and x_i is specified as $u_i = \delta_0 + \delta_1 y_{i,1} + \delta_2 x_i + \omega_i$ in order to integrate out the unobservable u_i from the joint density:

$$(1.30) \quad f(y_{i,2}, y_{i,3}, \dots, y_{i,T} | y_1, x_{i,t}, u_i) = \int f(y_{i,2}, y_{i,3}, \dots, y_{i,T} | y_{i,1}, x_{i,t}) j(u_i | y_{i,1}, x_{i,t}) du_i.$$

u_i is assumed to be normally distributed, $\omega_i \sim N(0, \sigma_\omega^2)$ and x_i is a vector of time variant explanatory variables in all time periods. Following Mundlak (1978), time averages of x_i , denoted as \bar{x}_i , can be used as x_i .⁷ $y_{i,t}$ given $(y_{i,t-1}, y_1, x_i, \omega_i)$ then turns out to follow a probit model

⁷ Mundlak (1978) suggests the correlation between u_i and the observed characteristics be controlled by assuming

$$u_i = \gamma \bar{x}_i + \eta_i$$

$$(1.31) \quad P(y_{it} = 1 | y_{i,t-1}, y_1, x_i, \omega_i) = \Phi(\alpha y_{i,t-1} + \beta x_{i,t} + \delta_0 + \delta_1 y_{i,1} + \delta_2 \bar{x}_i + \omega_i),$$

$$i = 1, 2, \dots, n; \quad t = 2, 3, \dots, T,$$

This is the Wooldridge Conditional Maximum Likelihood (CML) estimator. Parameter estimation can be done with standard random effects probit software.

1.4 Maximum Simulated Likelihood (MSL)

Both Heckman and Woodridge estimators are inevitably sensitive to the normality assumption on the density of u_i . If it is not correctly specified, the parameter estimation by maximum likelihood will be inconsistent. If the error term is serially correlated, the integral will be T dimensional; the integral will be analytically intractable and there will be no close-form expression for the likelihood function. In this case, maximum simulated likelihood (MSL) is warranted.

Unlike maximum likelihood (ML) that evaluates the probability of equation analytically, MSL simulates the probability. In the case of the Heckman approach, MSL is to approximate the probability by averaging the density $f(y_{i,2}, y_{i,3}, \dots, y_{i,T} | y_{i,1}, x_{i,t}, u_i)$ over many draws from the marginal distribution $h(u_i | x_i)$ while specifying a conditional density of the first observation $y_{i,1}$. There are a range of simulators that can compute the integral. The estimator summarized in this section is a random effect dynamic probit model with autocorrelated errors using the GHK

where $\eta_i \sim N(0, \sigma_\eta^2)$. This transforms equation (1.1) into

$$P(y_{it} = 1 | y_{i,t-1}, x_i, \eta_i) = \Phi(\alpha y_{i,t-1} + \beta x_{i,t} + \gamma \bar{x}_i + \eta_i), \quad i = 1, 2, \dots, n; \quad t = 2, 3, \dots, T,$$

by adding time averages of all time variant exogenous regressors. Consistent estimation of α requires independence between $y_{i,t-1}$ and η_i . If the true initial conditions are not observed in the sample, α will still be overestimated.

algorithm based on Halton draws.⁸ A Halton sequence⁹ is generated by dividing the unit interval into a multiple of p segments, where p is usually a prime, and arranging the breakpoints in a particular way. For example, if p is taken to be 3, the unit interval is first divided into three segments with breaks at $\frac{1}{3}$ and $\frac{2}{3}$. Then each of the three segments is further divided into three sub-segments, with first breakpoint of the first segment at $\frac{1}{9}$, first breakpoint of the second segment at $\frac{4}{9}$ and first breakpoint of the third segment at $\frac{7}{9}$. This is followed by the second breakpoints of the three segments: $\frac{2}{9}$, $\frac{5}{9}$ and $\frac{8}{9}$. This process goes on for the third and higher rounds of division in the same fashion giving breakpoints $\frac{1}{27}$, $\frac{10}{27}$, $\frac{19}{27}$, $\frac{4}{27}$, $\frac{13}{27}$ and so on.

The Halton sequence is then the set of all the breakpoints arranged in the order they have occurred: $(\frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{4}{9}, \frac{7}{9}, \frac{2}{9}, \frac{5}{9}, \frac{8}{9}, \frac{1}{27}, \frac{10}{27}, \frac{19}{27}, \frac{4}{27}, \frac{13}{27}, \dots)$. The Halton sequences are quasi-random numbers: their elements are selected systematically and two consecutive numbers are negatively correlated. Halton sequences provide more even coverage of the area of integration than do the same length of sequences of pseudo-random numbers, leading to smaller variance of the simulator.

Four dimensions of Halton sequences are created and the initial 20 elements of each sequence are discarded due to the high correlation among them. The initial period $y_{i,1}$ is specified as a function of all the exogenous variables of the first period.

⁸ Estimation can be done in Stata version 10 using redpacc command. It is an extension of the Heckman estimator.

⁹ See Train (2003) for a detailed discussion on Halton sequences.

The next chapter presents applications and empirical findings of the three estimators of dynamic panel data models summarized in this chapter. Estimation results are relatively consistent across estimators.

Chapter Two

This chapter examines the impact of the large price increases in cigarettes after the Master Settlement Agreement (MSA) on drinking behavior using data from the Panel Study of Income Dynamics (PSID). Dynamic panel data models are estimated. I address the unobserved heterogeneity and initial conditions issues using the estimation strategies summarized in chapter one, including system GMM, Wooldridge conditional maximum likelihood (CML) and simulated maximum likelihood (MSL). Alcohol consumption, drinking participation and heavy drinking participation (three or more drinks per day) are considered for the full sample, as well as for sub-samples stratified by age group and gender. Estimation results are relatively stable across estimators. I find that the cross-price effects of cigarettes on alcohol consumption are insignificant showing that averaging on all consumption levels, the number of drinks consumed per day is not affected by the increases in cigarette prices; and that the cross-price effects of cigarettes on drinking participation are mostly positive and significant, indicating drinking is an economic substitute for smoking; also, cigarette prices do not affect heavy drinking prevalence.

2.1 Introduction

Since the release of the Surgeon General's report on smoking and health in 1964, the health hazard caused by smoking has been more and more recognized. Smoking harms almost every organ of the body; and it remains the single largest preventable cause of death in the United States. Each year, there are 393,600 estimated deaths caused by direct smoking.¹⁰ Secondhand smoking is no less harmful. Secondhand smoke contains more toxic chemicals than does smoke inhaled from a cigarette. Currently, there are 49,400 estimated deaths of

¹⁰ Source: Center for Disease Control and Prevention (CDC). "Tobacco Use: Targeting the nation's leading killer." <http://www.cdc.gov/chronicdisease/resources/publications/AAG/osh.htm>. Accessed on July 11, 2011.

nonsmokers annually from secondhand smoking.¹¹ For half a century, the federal and state governments have been engaged in anti-smoking campaigns. Major policy instruments include widespread dissemination of information on the health consequences of cigarette smoking, ban of broadcast advertising of cigarettes, restrictions on smoking in public places and private workplaces, increased cigarette excise taxes and more (Chaloupka and Grossman 1996). Over the past decade or so, America has seen the historically largest increases in cigarette excise taxes. Since January 1 1997, with the exception of Florida, Mississippi, Missouri, North Dakota, South Carolina and Virginia, all other states and the District of Columbia have increased their taxes on cigarettes at least once. The following are the number of states as of 2008 that raised cigarette taxes by 25 cents or more per pack in nominal terms from the previous year: 3 states in 1997 and 1998, 1 in 1999, 2 in 2000, 7 in 2002, 18 in 2003, 8 in 2004, 9 in 2005 and 6 in 2006 through 2008. Among these, 13 states had one-time increases of a dollar or more at least once. As of 2008, New York has the highest state excise tax of \$2.75 per pack (Orzechowski and Walker 2008). Also over the past decade, an increasing number of states have passed stronger smoke-free air laws. As of March 2011, 27 states have enacted statewide bans on smoking in all enclosed public places, including private workplaces, restaurants and bars, whereas there was none prior to the year 2002.

Alcohol consumption, unlike cigarette smoking, imposes harm to the drinkers and others mainly when it reaches a certain amount. The common terms associated with the intensity of alcohol consumption, such as light drinking, moderate drinking and heavy drinking, are not standardized. Among physicians, there is an informal consensus as to what they mean by these terms. They operationally define light drinking as 1.2 drinks per day, moderate drinking as 2.2

¹¹ Source: [Facts about secondhand smoke](http://www.co.dakota.mn.us/HealthFamily/HealthyLiving/Smoking/ShSfacts.htm). <http://www.co.dakota.mn.us/HealthFamily/HealthyLiving/Smoking/ShSfacts.htm>. Accessed on July 15, 2011.

drinks per day, and heavy drinking as 3.5 drinks per day (Abel, Kruger and Friedl 1988). According to the Dietary Guidelines for Americans, drinking in moderation is defined as having no more than 1 drink per day for women and no more than 2 drinks per day for men.¹² The National Institute on Alcohol Abuse and Alcoholism (NIAAA) describes moderate drinking as consuming four drinks on any day with an average of 14 drinks per week for men; and three drinks in any one day and an average of seven drinks per week for women.¹³ Center for Disease Control and Prevention (CDC) defines heavy drinking as consuming an average of more than two drinks per day for men and consuming an average of more than one drink per day for women.¹⁴ In this paper, heavy drinking refers to having three or more drinks per day, and moderate drinking is having two drinks or less per day. Various studies in the medical literature suggest that moderate drinking is beneficial to heart health and longevity.¹⁵ It is the excessive alcohol consumption that causes serious health consequences such as cirrhosis of the liver and personal injuries such as from motor vehicle crashes, violence and crimes (Dave and Saffer 2008). Since the mid-1970s, measures have been taken to reduce alcohol-induced motor vehicle deaths, such as raising the minimal legal drinking age to 21 and lowering the legal limit of blood alcohol concentration (BAC) from 0.10% to 0.08%.

Although campaigns against smoking and drinking have been going on for decades, the adverse health and social effects related to cigarettes and alcohol are still major concerns of the public. Due to their addictive nature, the consumption of cigarettes and alcohol often affect each other. It would make a more effective policy to reduce the consumption of one good while

¹² United States Department of Agriculture and United States Department of Health and Human Services. In: *Dietary Guidelines for Americans*. Chapter 9 – Alcoholic Beverages. Washington, DC: US Government Printing Office; 2005, p. 43–46.

¹³ Source: “Alcohol: Problems and Solutions.” <http://www2.potsdam.edu/hansondj/AlcoholAndHealth.html>. Accessed on July 15, 2001.

¹⁴ Source: Center for Disease Control and Prevention. <http://www.cdc.gov/alcohol/faqs.htm#heavyDrinking>. Accessed on July 18, 2011.

¹⁵ See the discussion section for a brief review.

taking into account its effect on the consumption of the other. However, earlier studies have inconsistent findings regarding the complementarity or substitutability of the two goods. The next section summarizes the literature.

This paper takes advantage of the large cigarettes price hikes after the Master Settlement Agreement (MSA) to examine whether drinking is an economic complement to or substitute for cigarette smoking. In November 1998, 46 states and the four major tobacco companies reached the Master Settlement Agreement which stipulated that the tobacco companies pay the states \$206 billion over the next 25 years to compensate most states for Medicaid expenses for treating tobacco-related illnesses. The other four states (Florida, Mississippi, Texas and Minnesota) settled with the tobacco companies individually. The tobacco companies are financing these payments by increasing cigarette prices. As a result, the cigarette prices went up by 45 cents per pack¹⁶, or 19.5%, nationwide immediately after the settlement, and continued to rise over the following five years in many states due to substantial increases in their excise taxes on cigarettes.

In this chapter, I investigate how drinking behavior among adults has been affected by these unprecedented increases in cigarette prices and the changes in anti-smoking regulations (the full prices of smoking). I address the unobserved heterogeneity and initial conditions issues using the conditional maximum likelihood (CML) estimator. System GMM estimations with fixed effects are used as robustness checks. Since heavy drinking is more harmful and costly, heavy drinking participation and light drinking participation are considered separately. Results show that the cross-price effects of cigarettes on drinking participation are generally positive and significant, indicating drinking is an economic substitute for smoking; cigarette prices do not affect heavy drinking prevalence; there is some evidence that heavy drinking might be a complement to smoking for drinkers.

¹⁶ Adjusted to 1999 dollars.

The remainder of the chapter is organized as follows. The next section is a literature review. Section 3 presents a conceptual framework and the empirical model. Section 4 describes the data. Section 5 discusses the estimation strategies. Section 6 summarizes the results. Finally, section 7 concludes.

2.2 Background and Literature

Raising cigarette taxes has been one of the most effective means to prevent and reduce smoking. There is an extensive literature documenting the impact of changes in cigarette taxes or prices on the demand for cigarettes. Estimates of the price elasticity of demand for cigarettes from previous studies are between -0.3 and -0.5. Chaloupka and Warner (2000) provide a comprehensive literature review on demand for cigarettes.

Relatively few studies have focused on the cross-price effects of cigarettes on demand for alcohol. The following papers investigate the complementarity and substitutability of smoking and drinking. The results are mixed. Goel and Morey (1995) use US state level panel data from 1959 to 1982 to estimate demand equations for cigarettes and liquor allowing for cross-price effects. They find that cigarettes and liquor are substitutes in consumption. In an effort to address endogeneity in the structural models of interdependence between alcohol and cigarettes consumption, Dee (1999) utilizes within-state variation in cigarette taxes and the movement away from a minimum legal drinking age of 18 to conduct reduced form and instrumental variable estimations for separate drinking and smoking equations. Using state-level rates of drinking and smoking for teenagers, Dee finds that participation in drinking and smoking by teens are complements. Decker and Schwartz (2000) investigate cross-price effects in cigarettes and alcohol participation and consumption. Using a sample of nearly 500 thousand observations

drawn from the 1985-1993 Behavioral Risk Factor Surveillance System (BRFSS), the authors find that increases in the price of cigarettes increase drinking participation, suggesting that alcohol and cigarettes are substitutes, while increases in the price of alcohol decrease smoking participation, indicating that they are complements. Cameron and Williams (2001) use Australian individual level data for adults (NDSHS) to estimate probit models of the price responsiveness of participation in cannabis, alcohol and cigarette use. They regress participation indicators for each of the drugs (cannabis, alcohol and cigarette) on their full prices, interactions between the different prices and socioeconomic and demographic variables. They find some evidence that alcohol and cigarettes are complements. In particular, the effect of the price of cigarettes on alcohol participation is negative and strongly significant while the price of alcohol is positive but insignificant in the cigarette equation. In a similar study using more recent surveys of NDSHS, Zhao and Harris (2004) investigate the interrelations among participation decisions of marijuana, alcohol and tobacco. Extending Cameron and Williams (2001), the authors estimate a multivariate probit model to allow for correlations of an individual's unobserved heterogeneity across decisions about different drugs. They find stronger evidence for complementarity between alcohol and tobacco than that from Cameron and Williams (2001). Cross-price effects of tobacco and alcohol are both negative and significant at 5% in the alcohol and tobacco equations, respectively. Picone et al. (2004) focus on older persons above 50 years of age. Using the panel data Health and Retirement Survey (HRS) between 1992 and 2002, they estimate dynamic panel models using GMM estimators to analyze the effects of smoking bans and cigarette prices on alcohol consumption, and symmetrically, the effects of the 0.08 BAC per se laws and alcohol prices on cigarette consumption. The positive effect of cigarette prices on alcohol consumption suggests that alcohol and cigarettes are substitutes, but the negative effect

of smoking bans supports complementarity. Also estimating with GMM, Bask and Melkersson (2004) find that alcohol and cigarettes are complements using Swedish aggregate annual time series on sales volumes of cigarettes and alcohol for the period of 1955-1999. Markowitz and Tauras (2009) is the most recent paper in the literature that I am aware of. The authors estimate separate teens' participation demand equations for alcohol, cigarettes and marijuana with region fixed effects and individual fixed effects. What distinguishes their work from others is that they consider the influence of the prices of goods commonly bought by teenagers such as gasoline, clothing, fast food and entertainment on the demand for the three substances. Using the first four waves of NLSY97 from 1997 to 2001 matched with the price variables from ACCRA, they find evidence that cigarettes and alcohol are substitutes when the probability of alcohol use is examined, but no symmetric cross-price effects of alcohol when the probability of smoking is examined.

2.3 Conceptual Framework and Empirical Model

When modeling demand for one addictive consumer good, the rational addiction framework developed by Becker and Murphy (1988) is most widely used, in which both past and future consumption affect current consumption. Pacula (1998) and Kenkel et al. (2001) extend the rational addiction model to allow for two or more addictive goods for perfectly rational consumers.

The aim of this study is not to test if consumers are forward-looking in the rational addiction models. Rather, it is to assess whether the increases in the full price of cigarettes after the MSA affect alcohol use when allowing for reinforcement effects. Reinforcement of an addictive good means that past consumption increases current consumption by raising the

marginal utility of current consumption. Due to the addictive nature of alcohol and cigarettes, past consumption of alcohol will reinforce its current consumption, and past consumption of cigarettes may also reinforce current consumption of alcohol. Therefore, the model is a myopic model of two addictive goods where future consumption does not influence current decisions.

Following Picone et al. (2004), the model starts with an individual's utility in period t as a function of the consumption of alcohol (A_t), cigarettes (C_t) and the composite good (G_t) given consumption of alcohol and cigarettes from last period (A_{t-1}) and (C_{t-1}). The individual maximizes utility:

$$U = u(A_t, C_t, G_t | A_{t-1}, C_{t-1})$$

subject to a budget constraint:

$$Y_t = p_t^A A_t + p_t^C C_t + p_t^G G_t$$

where Y_t is income in period t , p_t^A , p_t^C and p_t^G are prices for alcohol, cigarettes and the composite good, respectively. Prices of alcohol and cigarettes are the full prices that include monetary prices and policies that regulate the consumption of alcohol and cigarettes.

Maximizing utility subject to the budget constraint yields alcohol demand function for the individual. The dependent variable can be either amount of alcohol consumed or a dichotomous indicator for drinking. The following empirical model for the probability of drinking for individual i in period t will be estimated

$$(2.1) \quad d_{i,t} = \alpha_1 d_{i,t-1} + \alpha_2 s_{i,t-1} + \beta_1 p_{i,t}^d + \beta_2 p_{i,t}^s + \beta_3 pol_{i,t}^d + \beta_4 pol_{i,t}^s + \beta_5 x_{i,t} + \varepsilon_{i,t}$$

where the error term $\varepsilon_{i,t}$ consists of the unobserved individual fixed effects u_i and the idiosyncratic disturbances $v_{i,t}$ ($\varepsilon_{i,t} = u_i + v_{i,t}$). $d_{i,t}$ is the amount of alcohol consumed or an indicator for drinking participation, $d_{i,t-1}$ and $s_{i,t-1}$ are drinking and smoking statuses for the

previous period¹⁷, $p_{i,t}^d$ and $p_{i,t}^s$ are the money prices for alcohol and cigarettes, $pol_{i,t}^d$ and $pol_{i,t}^s$ are the policy variables for drinking and smoking, $x_{i,t}$ is a vector of exogenous social-demographic variables including family income, age, gender, race, ethnicity, education, employment and marital status. α_1 measures the reinforcement effect of past drinking behavior on current drinking, and it is expected to be positive. α_2 indicates whether smoking is a substitute for or a complement to drinking; a priori, its sign is unknown. β_1 and β_3 , the coefficients on the own-price and policy of alcohol, should be negative based on the law of demand. Interest lies in the signs and significance of β_2 and β_4 , the coefficients on the cross-price and cross-policy of alcohol. They can be negative or positive depending on whether drinking is a complement to smoking or a substitute for smoking.

Estimation of equation (2.1) requires longitudinal data. Difficulty arises in obtaining unbiased and consistent estimates of the state dependence due to the correlation between the time invariant unobserved individual heterogeneity u_i and the lagged drinking and smoking behaviors. Serial correlation in the idiosyncratic disturbances $v_{i,t}$ also complicates the estimation. Even when coefficients on lagged dependent variables are not of direct interest, estimating them consistently may be crucial for recovering consistent estimates of other parameters (Bond 2002).

2.4 Data

The data are from the 1999, 2001, 2003, 2005 and 2007 waves of the Panel Study of Income Dynamics (PSID). Conducted by the Survey Research Center, Institute for Social

¹⁷ A specification using cigarette consumption instead of smoking participation is also estimated.

Research, University of Michigan, the PSID is a longitudinal study of a nationally representative sample of U.S. individuals and the family units in which they reside. The initial wave of the PSID was administered in 1968. Follow-up interviews were conducted annually until 1996 and biennially thereafter. The health behavior measures such as alcohol and cigarette consumption have been collected since 1999.

Dependent variables

The research objective is to find out whether drinking behavior is significantly affected by changes in cigarette prices. Three dependent variables are used: a measure of average number of drinks per day the individual consumes, drinking participation and heavy drinking participations. In the first three waves, respondents were asked “On average, do you have less than one drink a day, one or two drinks in a day, three to four drinks a day, or five or more drinks a day?” In the last two waves, they were asked “On average, how often did you have any alcohol to drink, less than one a month, about once a month, several times a month, about once a week, several times a week, or every day?” and “On the days you drank, about how many drinks did you have?” For the first three waves, the alcohol consumption variable is approximated by assigning a zero to the category of having no or less than one drink per day, 1.5 to having one or two drinks per day, 3.5 to having three to four drinks per day and 5 to having 5 or more drinks per day. For the years 2005 and 2007, this variable is generated by dividing the approximate number of drinks consumed in a month by 30 days. Two dichotomous indicators for drinking participation are constructed. One is a dummy variable equal to one to indicate that the respondent drank any amount of alcohol and equal to zero otherwise. The other is a dummy variable that indicates whether the respondent drank more than two drinks per day. For the years

2005 and 2007, this variable is created based on the daily alcohol consumption variable mentioned above.

Independent variables

The model is a myopic addiction model and assumes past alcohol consumption reinforces current alcohol consumption. A binary indicator of drinking participation in the previous wave is included. A positive sign is expected on this variable. The model also tests whether past cigarette consumption affects current drinking behavior. Variables for past smoking participation and smoking intensity are created based on the responses to the questions “Do you smoke cigarettes?” and “On the average, how many cigarettes per day do you usually smoke?” They are used in different specifications.

The PSID provides a wide range of demographic and socioeconomic variables. The estimation of the model controls for age, sex, race, ethnicity, educational attainment, family income, household size, marital and employment status. Three categories of race are used: white, black and other races. White is the omitted group. Four categories of educational attainment are created: college, some college, high school and less than high school. College is defined as 16 or more years of schooling; some college is defined as 13 to 15 years of schooling; high school is 12 years of schooling and less than high school is 11 or fewer years of schooling. College is chosen to be the omitted group. Family income is the total family income in thousands of dollars deflated to 1999 dollars from the previous year of interview. Household size is the number of persons living in the same family unit. For marital status, a dichotomous indicator for marriage is constructed. Similarly for employment status, a dichotomous indicator

is used which is equal to one if the individual is working now or only temporarily laid off or on sick/maternity leave, and zero otherwise.

The key variables of the model are the costs of consuming alcohol and cigarettes. The costs consist of monetary costs and non-monetary costs. State level prices of a six-pack of beer constructed from the American Chamber of Commerce Researchers' Association (ACCRA) are used as proxies for the money prices of alcohol: the quarterly beer prices of 250 to 300 localities are adjusted to ACCRA cost of living indices and are averaged for each state by year. Cigarette prices come from the state level weighted average prices per pack in the Tax Burden on Tobacco (Orzechowski and Walker 2008). All prices are deflated to 1999 dollars. Non-monetary cost refers to policies and regulations that increase the degree of inconvenience for consuming alcohol and cigarettes. The model includes two such variables: a smoking ban index and a dummy variable indicating whether a state has an effective 0.08 BAC law for the year of interview. Under the 0.08 BAC laws, it is illegal to operate a motor vehicle if the driver has a blood alcohol concentration of 0.08% or higher. During the sample period from 2001 to 2005, 23 states lowered the legal BAC limit from 0.10 to 0.08 while the other 27 states had done so prior to 2001. This leaves some variation in the timing of enactment of the laws. The smoking ban index is constructed based on the smoking restriction decisions of the smoke-free air laws in the following 12 locations: Government worksites, private worksites, child care centers, health care facilities, restaurants, recreational facilities, cultural facilities, public transit, shopping malls, public schools, private schools, and free standing bars. Smoking restrictions for some locations are coded 0, 1, 2 and 3, with 0 being no restrictions against smoking and 3 being smoking banned at all times; other locations are coded from 1 to 5 in the similar fashion. The smoking ban index for each state of each year is the sum of the numerical codes. The price and policy variables are

merged with the PSID data based on the respondent's state of residence and the year of interview. The law of demand predicts negative signs on beer prices and 0.08BAC. The signs of cigarette prices and the smoking ban index indicate whether alcohol and cigarettes are complements or substitutes. If they are substitutes, the model predicts positive signs. Tables 1 and 2 show the summary statistics of the variables.

2.5.1 Estimation of Equation (2.1)

Equation (2.1) will be first estimated as a linear dynamic panel data model with fixed effects using the two-step system Generalized Method of Moment (GMM).¹⁸ Alcohol consumption is considered using GMM.

Since drinking and smoking may reinforce each other, $d_{i,t-1}$ and $s_{i,t-1}$ are both positively correlated with time invariant preferences for alcohol, $u_{i,t}$. Therefore, treating all other variables as strictly exogenous, OLS estimation of equation (2.1) will yield upward biased and inconsistent estimates. Fixed effects using within group transformation will have downward biased and inconsistent estimates.

A system of the first differenced equation and the original equation in levels are estimated

(2.2)

$$\begin{aligned}\Delta d_{i,t} &= \alpha_1 \Delta d_{i,t-1} + \alpha_2 \Delta s_{i,t-1} + \beta_1 \Delta p_{i,t}^d + \beta_2 \Delta p_{i,t}^s + \beta_3 \Delta pol_{i,t}^d + \beta_4 \Delta pol_{i,t}^s + \beta_5 \Delta x_{i,t} + \Delta v_{i,t} \\ d_{i,t} &= \alpha_1 d_{i,t-1} + \alpha_2 s_{i,t-1} + \beta_1 p_{i,t}^d + \beta_2 p_{i,t}^s + \beta_3 pol_{i,t}^d + \beta_4 pol_{i,t}^s + \beta_5 x_{i,t} + \varepsilon_{i,t}\end{aligned}$$

$d_{i,t}$ is the amount of alcohol consumed; $d_{i,t-1}$ and $s_{i,t-1}$ are predetermined variables: past participation in drinking $d_{i,t-1}$ is not correlated with current shocks to drinking $v_{i,t}$, but are

¹⁸ GMM estimations are done using the command xtabond2 in Stata version 10.

affected by the past ones; past smoking participation/consumption $s_{i,t-1}$ may also be affected by the past disturbances in drinking. System GMM uses the lagged levels of the predetermined variables as instruments for their first-differenced form and uses the first differenced predetermined variables (now exogenous to the fixed effects) to instrument them in levels. Let $\Delta X_{i,t}$ be a vector of all exogenous variables in first differences

$\Delta X_{i,t} = (\Delta p_{i,t}^d, \Delta p_{i,t}^s, \Delta pol_{i,t}^d, \Delta pol_{i,t}^s, \Delta x_{i,t})$. I use $X_{i,t}$ to instrument $\Delta X_{i,t}$ in the difference equation and to instrument themselves in the level equation. One period lagged price and policy variables can serve as the additional instruments for the predetermined variables.¹⁹ Define

$LP_{i,t} = (p_{i,t-1}^d, p_{i,t-1}^s, pol_{i,t-1}^d, pol_{i,t-1}^s)$. Instrument matrices are given by²⁰

$$(2.3) \quad Z_{di} = \begin{pmatrix} d_{i1}s_{i1} & 0 & 0 & X_{i3}\Delta LP_{i,3} \\ 0 & d_{i1}d_{i2}s_{i1}s_{i2} & 0 & X_{i4}\Delta LP_{i,4} \\ 0 & 0 & d_{i1}d_{i2}d_{i3}s_{i1}s_{i2}s_{i3} & X_{i5}\Delta LP_{i,5} \end{pmatrix}$$

$$(2.4) \quad Z_{li} = \begin{pmatrix} \Delta d_{i1}\Delta s_{i1} & 0 & 0 & X_{i3}LP_{i,3} \\ 0 & \Delta d_{i1}\Delta d_{i2}\Delta s_{i1}\Delta s_{i2} & 0 & X_{i4}LP_{i,4} \\ 0 & 0 & \Delta d_{i1}\Delta d_{i2}\Delta d_{i3}\Delta s_{i1}\Delta s_{i2}\Delta s_{i3} & X_{i5}LP_{i,5} \end{pmatrix}$$

$$(2.5) \quad Z_{si} = \begin{pmatrix} Z_{di} & 0 \\ 0 & Z_{li} \end{pmatrix}$$

The moment conditions are²¹

$$(2.6) \quad E(Z'_s Q) = 0$$

¹⁹ Stata also allows excluded IV style instruments in addition to the GMM style instruments.

²⁰ Z_{di} is the instrument matrix for the differenced equation for each individual; Z_{li} is the instrument matrix for the level equation for each individual; Z_{si} is the instrument matrix for the system of equations for each individual.

²¹ $Z_s, \Delta U_i$ and U_i are stacked matrices across individuals.

$$\text{where } Q = \begin{pmatrix} \Delta U \\ U \end{pmatrix}, \Delta U_i = \begin{bmatrix} \Delta v_{i3} \\ \Delta v_{i4} \\ \vdots \\ \Delta v_{iT} \end{bmatrix} \text{ and } U_i = \begin{bmatrix} v_{i3} \\ v_{i4} \\ \vdots \\ v_{iT} \end{bmatrix}$$

Windmeijer finite-sample corrected standard errors are reported along with the parameter estimates. The validity of the instruments depends on serially uncorrelated idiosyncratic errors $v_{i,t}$. Tests of this presumption are applied to the differenced residuals $\Delta v_{i,t}$. Estimates using matrix (2.5) will be inconsistent if $\Delta v_{i,t}$ exhibit second-order autocorrelation which indicates autocorrelation in $v_{i,t}$. Stata reports the Sargan/Hansen overidentification tests which test the null hypothesis of exogeneity of the instruments as a group. The Sargan's statistic, which is consistent only under homoskedasticity, is a special form of the Hansen's statistic. Since homoskedasticity is not assumed in two-step GMM, the Hansen's statistic is superior. P-values associated with the Hansen's statistics are reported. Note that the power of the Sargan/Hansen test becomes weaker as the number of instruments increases because the more instruments, the harder it is to minimize all the quadratic distances of the moment conditions towards zero (Roodman, 2006). Besides testing the validity of the whole instrument set, it is often helpful to check the validity of a subset of instruments by conducting the Difference-in-Hansen test. For the system GMM estimations in this paper, the exogeneity of the instruments for the level equations are tested. The p-values associated with these Difference-in-Hansen statistics are reported. Not rejecting the Difference-in-Hansen test also sheds light on the validity of the initial condition assumption specified in 1.15.

Next, the Wooldridge CML estimator will be applied to the model equation as

$$(2.7) P(d_{i,t} = 1) = \Phi(\alpha_1 d_{i,t-1} + \alpha_2 s_{i,t-1} + \beta_1 p_{i,t}^d + \beta_2 p_{i,t}^s + \beta_3 pol_{i,t}^d + \beta_4 pol_{i,t}^s + \beta_5 x_{i,t} + \delta_1 d_{i,1} + \delta_2 s_{i,1} + \delta_3 \bar{x}_i + \omega_i)$$

where $i = 1, 2, \dots, n$, $t = 2, 3, \dots, T$, $d_{i,t}$ is an indicator for drinking participation, $d_{i,1}$ is the initial period drinking participation status, $s_{i,1}$ is the initial period smoking participation/consumption, and \bar{x}_i is a vector of the time averages of age, educational attainment, family income, household size, marital status and employment status for each individual. Results from pure random effects are also reported as a comparison.

Finally, the random effect dynamic probit model with autocorrelated errors based on 100 Halton draws are implemented in Stata using the `redp` command. Four dimensions of Halton sequences are created and the initial 20 elements of each sequence are discarded due to the high correlation among them. The initial period $y_{i,1}$ is specified as a function of all the exogenous variables of the first period.

2.5.2 Estimation Results

This section presents estimation results from the system GMM, CML and MSL estimators. Estimates for alcohol consumption using system GMM are reported first. Estimates for drinking participation and heavy drinking participation from CML and MSL are then presented. Separate regressions are done for the full sample, as well as for sub-samples stratified by age group and gender.²² Estimates using lagged cigarette consumption in lieu of lagged smoking participation are also reported.

²² Unbalanced panels are used. The sub-sample of age of 50 or below, for example, contains the observations aged 50 or younger between the year 2001 and 2007. If a person reached age 50 in 2005, he or she is observed three times. The `redp` Stata command, however, requires the dependent variable of each individual to be observed for all waves. Therefore, estimates from MSL for the sub-samples are based on different sample sizes.

2.5.3 Alcohol Consumption Using GMM

Table 3 contains OLS and FE estimates for the alcohol consumption equation. OLS estimates show that the lagged alcohol consumptions are positive and significant at the 1% level, whereas estimates from the fixed effects models are negative and significant at the 1% level. To take into account of any time-specific common trends, year fixed effects are included in all analyses. However, state fixed effects are not used to control for the unobserved state sentiment. During the seven year sample period, there is not much variation in beer prices within each state. Adding state fixed effects takes away most of the variation. Following Dave and Saffer (2008), regressing beer prices on state and year fixed effects yields an R-squared value of 0.99, meaning 99% of the within-state variation in beer prices are taken away and only 1% is left to explain changes in drinking behavior. This results in unstable parameter estimates for alcohol prices and inflated standard errors. Thus, it is preferable to control for the unobserved state sentiment only through state level policies and regulations. The discussion of all results focuses on estimations with year fixed effects.

GMM estimations for the alcohol consumption equation are reported in Tables 4 and 5. The parameter estimates for lagged alcohol consumption range from 0.11 to 0.2, lower than those from OLS and higher than FE. This coincides with the expectation that OLS overestimates the state dependence while FE underestimates it. In Table 4, cigarette prices and smoking ban indices exhibit positive and significant effects on alcohol consumption only for females. In Table 5, lagged smoking participation is replaced with lagged cigarette consumption. Smoking ban indices continue to show positive and significant effects. However in both specifications, this sub-sample fails the Hansen test of which the null hypothesis is that all instruments as a

group are exogenous. Therefore, there is no evidence as to whether alcohol consumption is an economic complement to or substitute for smoking.

2.5.4 Drinking Participation

CML

Table 6 contains OLS, FE and RE estimates for drinking participation. Similar to the results in Table 1, lagged drinking participation from OLS is positive and significant at the 1% level while the FE shows negative and significant effects on drinking prevalence at the 1% level. The RE estimates are very close to OLS in magnitudes and significance.

CML estimations for drinking participation equations are reported in Tables 7 and 8. The marginal effects of lagged drinking participation are again between OLS/RE and FE. In Table 7, lagged smoking participation exhibits insignificant effects on current drinking participation for all sample specifications. The own-price and policy variables have negative effects in all cases. For the full sample, a dollar increase in a six-pack of beers reduces the probability of drinking by two percentage points and the effect is significant at the 10% level; the 0.08 BAC law is associated with a 4.6 percentage point decrease in the probability of drinking and the effect is significant at the 1% level, all else equal. For the sub-samples, beer price is significant for younger adults and females; 0.08 BAC is statistically significant for people of all ages and for females. The cross-price and cross-policy variables all have positive signs and are significant at the 5% or 1% level except for the age group of 50 or younger. For the full sample, a dollar increase in the price of a 20-pack of cigarettes increases the probability of drinking by 4.2 percentage points *ceteris paribus*; the effect is significant at the 1% level. Similarly, a one unit increase in smoking ban index increases the probability of drinking by 0.08 percentage point and

the effect is significant at the 5% level. The Smoking ban index is also significant at the 5% level for females. Except for individuals younger than 50, there is sufficient evidence that alcohol is an economic substitute for cigarettes.

Drinking prevalence declines with age. For the full sample, one year increase in age is associated with 0.8 percentage point decrease in the probability of drinking participation holding all else constant; the effect is significant at the 1% level. For ages above 50, there is a 2.1 percentage point decrease in the probability of drinking; the effect is significant at the 1% level. For males, its effect is 1 percentage point.

Males participate more in drinking than do females. Marital status affects drinking prevalence; married individuals tend to drink significantly less than the unmarried. Relative to the unemployed, employed individuals have higher drinking participation possibly because they are more likely to be social drinkers. This is also true for males only.

Due to the addictive nature of cigarettes, the increases in cigarette prices and the more stringent smoke-free air laws after the MSA may induce an individual to reduce cigarette consumption rather than quit smoking. To see how current drinking participation is affected by cigarette prices and smoke-free air laws considering the marginal changes in the quantity of cigarettes smoked in the past period, lagged smoking participation is replaced with lagged smoking intensity as in the case of GMM. Table 8 reports marginal effects from CML of this specification for all samples with year fixed effects. Results are nearly identical to those from the specification with lagged smoking participation in Table 7.

MSL

Table 9 reports selected parameter estimates from MSL with serially correlated errors. Panel A uses lagged smoking participation as an independent variable and panel B uses lagged

cigarette consumption. In both panels, the coefficients on lagged drinking are all close to those from RE models across columns. Drinking prevalence responds negatively to alcohol's own prices and policy; 0.08 BAC laws have significant effects in all columns. Cigarette prices are all positive and very significant. For females, smoking bans is also positive and significant. Comparing with CML and RE, it is possible that the state dependence is overestimated here. Yet, the significance level of cigarette prices and smoking bans is consistent with CML.

2.5.5 Heavy Drinking Participation

CML

In section 2.6.2, drinking has been consistently found a substitute for smoking for the full sample, older adults, males and females. I did not find such evidence for individuals 50 years of age or younger. This subsection reports how heavy drinking participation is affected by the changes in the cost of smoking. Table 10 shows the marginal effects of cigarette prices and smoking ban indices on heavy drinking participation. Panel A contains the estimation results using lagged smoking participation as an independent variable, and panel B uses number of cigarettes smoked. The cost of smoking has no significant effects on heavy drinking prevalence indicating that heavy drinking is not a substitute for smoking.

MSL

Table 11 reports selected parameter estimates for heavy drinking using MSL. Panel A contains the estimation results using lagged smoking participation as an independent variable, and panel B uses number of cigarettes smoked. Consistent with the CML estimates, heavy drinking prevalence is not affected by the full price of smoking.

2.6 Discussion

The insignificant effects of cross-price (full price) of cigarettes on alcohol consumption using GMM indicate that averaging on all consumption levels, the number of drinks consumed per day is not affected by the increases in cigarette prices. However, the CML and MSL estimates suggest that in response to the substantial increases in cigarette prices, individuals substitute away from cigarettes towards alcohol to the extent of a couple of drinks per day. The cross-price (full price) effects of cigarettes on drinking participation are generally positive and significant, but not significant on heavy drinking participation. The estimators used in this study are consistent and results are relatively stable across estimators.

One negative externality of the policy of preventing and reducing smoking by raising the full price of cigarettes could be increased alcohol-related harms. This research suggests otherwise. As other research has shown, light or moderate alcohol use has little negative social and health effects. It is the overuse and misuse that need to be deterred. Two alcoholic drinks generally raise the blood alcohol concentration to 0.05% for a 160-pound male, and to 0.08% for a female who weighs 120 pounds. As for health, a number of studies in the medical literature suggest that moderate drinking benefits heart health and longevity.²³

Rimm et al. (1991) find evidence that moderate alcohol consumption reduces the risk of coronary heart disease among men. Manson et al. (1996) conclude about coronary heart disease that consumption of one or two drinks of beer, wine, or liquor per day corresponds to a reduction in risk of approximately 20-40%. According to Sesso et al. (2001), among the 18,455 males in their study, those originally consuming one drink per week or less who increased their consumption to six drinks per week or less has a 29% reduction in Coronary Vascular Disease

²³ Moderation in these studies is generally defined as one to three drinks per day. It also depends on the body weight of the individual.

(CVD) risk compared to those who did not increase their consumption. Men originally consuming 1-6 drinks per week who increased their consumption moderately has a 15% decrease in CVD risk compared to those who made no change.

Similar research has been done on women. According to Dorn et al. (2007), women who drank moderately had a significantly lower risk of heart attack than abstainers, and the benefits were greatest in women who had a drink daily. Medical researchers at the State University of New York's University at Buffalo report that women who regularly consume an alcoholic drink or two have a significantly lower risk of having a non-fatal heart attack than women who are life-time abstainers.²⁴

Similar evidence has been found for the elderly. Simons et al. (2000) report that moderate alcohol intake appears to be correlated with less cardiovascular deaths in men 60-74 years and in all elderly women.

Moderate drinking is also associated with improved longevity. A study of more than 40,000 people by the Cancer Research Center in Honolulu has found that male moderate drinkers had a reduced risk for all-cause mortality of 14 to 22 percent than nondrinkers after controlling for confounding factors. For women overall, all-cause mortality rates were reduced by 18 percent (Maskarinec et al. 1998). Another major medical research study on 89,299 men in the Physicians' Health Study over a period of five and one-half years has found greater longevity among men who drink moderately than among those who either abstain or drink heavily (Gaziano *et al.* 2000). Italian researchers examined the relationship between alcohol

²⁴ State University of New York University at Buffalo. Moderate Drinking Lowers Women's Risk of Heart Attack. Press release, May 23, 2007.

consumption and long-term survival in 1,536 men aged 45-65. The results show that about two years of life are gained by moderate drinkers in comparison with occasional (those who drank less than 1 drink per week) and heavy drinkers (Farchi et al. 2000). Maraldi et al. (2006) examined 2,487 adults aged 70-79 years without heart disease at the beginning of the study. They were followed for an average period of over five and a half years. The results show that the risks of cardiac events (myocardial infarction, angina, or heart failure) and of all-cause mortality were significantly lower in light to moderate drinkers than in abstainers or occasional drinkers.

In conclusion, the findings of this research suggest that light or moderate alcohol use (two drinks or less per day) is a substitute for cigarette consumption, while heavy drinking is not. Increases in cigarette prices should not impose negative alcohol-related health or social effects.

Table 1. Alcohol and Cigarette Use Descriptive Statistics

Variables	Definition	Entire Sample	Ages 50 and Younger	Ages 51 and Older	Males	Females
Alcohol Consumption	Average number of drinks consumed per day	0.283 (0.830)	0.282 (0.844)	0.284 (0.803)	0.452 (1.068)	0.155 (0.556)
Drinking Participation	Dichotomous indicator for current drinkers	0.568 (0.495)	0.601 (0.490)	0.509 (0.500)	0.656 (0.475)	0.502 (0.500)
Heavy Drinking Participation	Dichotomous indicator for current heavy drinkers	0.031 (0.173)	0.031 (0.172)	0.031 (0.173)	0.054 (0.227)	0.013 (0.113)
Smoking Participation	Dichotomous indicator for current smokers	0.194 (0.395)	0.224 (0.417)	0.139 (0.346)	0.213 (0.409)	0.180 (0.384)
Cigarettes	Average number of cigarettes currently smoked per day	2.80 (7.21)	3.12 (7.36)	2.23 (6.89)	3.42 (8.38)	2.33 (6.14)
Observations		40395	25997	14398	17428	22967

This table reports means of the listed variables. Standard deviations are in parentheses. Numbers of observations listed are the maximum numbers in each column.

Some variables may have fewer observations due to non-response.

Table 2. Independent Variables Descriptive Statistics

Variables	Definition	Mean	Standard Deviation
Beer Price	Price of a six-pack of beer adjusted by consumer price index to 1999 dollars	6.08	1.10
Cigarette Price	Price of a pack of 20 cigarettes adjusted by consumer price index to 1999 dollars	3.47	0.53
0.08 BAC	Dichotomous indicator for having an effective 0.08 BAC law	0.77	0.42
Smoking Ban Index	Numerical values created from the smoking restriction of the smoke-free air laws	18.25	13.64
Age	Age of respondent	46.91	14.17
Male	Dichotomous indicator for male	0.43	0.50
White	Dichotomous indicator for white	0.63	0.48
Black	Dichotomous indicator for black	0.26	0.44
Other Race	Dichotomous indicator for race other than white or black	0.12	0.32
Hispanic	Dichotomous indicator for Hispanic	0.04	0.20
College	Dichotomous indicator for having college degree or higher	0.38	0.48
Some College	Dichotomous indicator for having some college	0.14	0.35
High School	Dichotomous indicator for having high school diploma or equivalent	0.29	0.45
Less than High School	Dichotomous indicator for not having completed high school	0.19	0.39
Family Income	Total family income of previous year in thousands of dollars discounted 1999 dollars	63.70	85.01
Household Size	Number of persons residing in the household	2.98	1.45
Married	Dichotomous indicator for being married	0.73	0.44
Employed	Dichotomous indicator for employed	0.70	0.46

Table 3. OLS and FE Estimates for Alcohol Consumption

Independent Variable	Aged 50 and younger				
	Full sample	Aged 50 and younger	Aged 51 and older	Males	Females
Panel A: OLS	(1)	(2)	(3)	(4)	(5)
Lagged alcohol consumption	0.586*** (0.028)	0.544*** (0.041)	0.650*** (0.022)	0.638*** (0.020)	0.463*** (0.062)
Lagged smoking participation	0.118*** (0.015)	0.125*** (0.019)	0.096*** (0.027)	0.168*** (0.027)	0.084*** (0.017)
Cigarette price	0.012 (0.011)	0.021 (0.017)	0.001 (0.013)	0.013 (0.017)	0.016 (0.014)
Smoking ban index	0.000 (0.0004)	-0.0004 (0.0005)	0.0005 (0.0007)	-0.0005 (0.0009)	0.0006 (0.0004)
Beer price	-0.008 (0.008)	-0.007 (0.010)	-0.010 (0.013)	-0.009 (0.015)	-0.006 (0.007)
0.08 BAC	-0.017 (0.014)	-0.027 (0.018)	-0.003 (0.021)	-0.030 (0.026)	-0.006 (0.014)
Panel B: FE					
Lagged alcohol consumption	-0.100** (0.041)	-0.130** (0.058)	-0.097*** (0.035)	-0.050 (0.032)	-0.204** (0.085)
Lagged cigarette consumption	-0.045* (0.025)	-0.053* (0.031)	-0.004 (0.036)	-0.028 (0.042)	-0.053* (0.032)
Cigarette price	0.003 (0.018)	0.018 (0.026)	-0.023 (0.026)	0.012 (0.035)	-0.005 (0.018)
Smoking ban index	0.0005 (0.0006)	0.0003 (0.0007)	0.001 (0.0001)	0.0008 (0.001)	0.0002 (0.0005)
Beer price	-0.023 (0.017)	-0.005 (0.021)	-0.038 (0.032)	-0.024 (0.032)	0.024 (0.016)
0.08 BAC	0.039 (0.017)	0.051** (0.025)	0.008 (0.023)	0.054 (0.035)	0.028 (0.015)
Number of observations	26554	16072	10482	11456	15098

Panel A uses lagged smoking participation as an independent variable;

Panel B uses lagged cigarette consumption.

Other control variables are omitted in the table.

All regressions include year fixed effects.

Standard errors are in parentheses.

*significant at 10% level, **significant at 5% level, ***significant at 1% level

Table 4. GMM Estimates for Alcohol Consumption with Lagged Smoking Participation

Independent Variable	Full sample (1)	Aged 50 and younger (2)	Aged 51 and older (3)	Males (4)	Females (5)
Lagged alcohol consumption	0.127*** (0.032)	0.243*** (0.046)	0.109** (0.043)	0.125*** (0.043)	0.125*** (0.048)
Lagged smoking participation	-0.078 (0.060)	0.009 (0.069)	-0.209** (0.099)	0.018 (0.128)	-0.124** (0.057)
Cigarette price	-0.037 (0.041)	0.032 (0.054)	-0.007 (0.067)	-0.017 (0.078)	0.054* (0.032)
Smoking ban index	0.0007 (0.0009)	0.002 (0.001)	-0.001 (0.001)	-0.0003 (0.002)	0.001** (0.0007)
Beer price	-0.0002 (0.035)	0.071** (0.034)	-0.008 (0.065)	0.075 (0.077)	0.009 (0.022)
0.08 BAC	0.008 (0.096)	0.026* (0.114)	-0.057 (0.154)	0.165 (0.178)	-0.056 (0.076)
Age	-0.047* (0.027)	0.021** (0.008)	-0.003 (0.019)	-0.055 (0.046)	-0.008 (0.011)
Male	1.001 (0.652)	0.254 (0.415)	2.381* (1.376)	-	-
Black	-0.219 (0.392)	-0.288 (0.326)	0.483 (0.870)	0.147 (1.133)	0.140 (0.201)
Other race	-0.276 (0.369)	-0.520 (0.332)	1.086 (0.949)	-3.399 (6.415)	-0.132 (0.210)
Hispanic	1.093 (1.376)	1.056 (0.837)	-5.657 (3.572)	-4.614 (7.377)	0.117 (0.662)
Some college	-0.121 (0.583)	0.217 (0.464)	-3.545** (1.649)	-4.780* (2.763)	0.215 (0.331)
High school	0.116 (0.459)	0.185 (0.344)	-1.988* (1.121)	-1.291 (2.685)	0.041 (0.224)
Less than high school	-0.011 (0.366)	0.048 (0.291)	-2.059** (1.023)	-2.389 (2.543)	-0.301*** (0.192)
Family income	0.0003**** (0.0001)	0.00005 (0.0002)	0.0005* (0.0003)	0.0002 (0.0001)	0.0006*** (0.00002)
Household size	0.024 (0.021)	0.028 (0.020)	-0.018 (0.044)	0.040 (0.040)	0.030 (0.020)
Married	0.005 (0.089)	0.042 (0.095)	0.127 (0.198)	0.061 (0.132)	0.084 (0.057)
Employed	0.031 (0.025)	0.033 (0.032)	-0.023 (0.040)	0.058 (0.045)	0.005 (0.024)
Number of observations	26554	16072	10482	11456	15098
Number of instruments	42	42	42	41	41
p-value AR(1) test	0.000	0.000	0.000	0.000	0.018
p-value AR(2) test	0.490	0.055	0.482	0.894	0.321
p-value Hansen test	0.019	0.001	0.061	0.713	0.013
p-value Difference-in-Hansen test	0.110	0.001	0.862	0.144	0.871

Windmeijer finite-sample corrected standard errors are in parentheses.

All regressions include year fixed effects.

*significant at 10% level, **significant at 5% level, ***significant at 1% level

Table 5. GMM Estimates for Alcohol Consumption with Lagged Cigarette Consumption

Independent Variable	Full sample (1)	Aged 50 and younger (2)	Aged 51 and older (3)	Males (4)	Females (5)
Lagged alcohol consumption	0.133*** (0.034)	0.245*** (0.047)	0.114** (0.045)	0.132*** (0.042)	0.116** (0.048)
Lagged cigarette consumption	-0.001 (0.004)	0.004 (0.005)	-0.008* (0.005)	-0.0001 (0.007)	-0.009*** (0.003)
Cigarette price	0.036 (0.042)	0.021 (0.052)	-0.010 (0.066)	-0.003 (0.077)	0.051 (0.032)
Smoking ban index	0.0003 (0.0009)	0.001 (0.001)	-0.001 (0.001)	-0.0003 (0.002)	0.001** (0.0007)
Beer price	-0.005 (0.036)	0.060* (0.109)	-0.003 (0.061)	0.070 (0.076)	0.003 (0.023)
0.08 BAC	0.031 (0.096)	0.060* (0.109)	-0.054 (0.163)	0.135 (0.176)	-0.026 (0.077)
Age	-0.050* (0.029)	0.021** (0.008)	0.000 (0.017)	-0.061 (0.054)	-0.018 (0.013)
Male	0.843 (0.656)	-0.091 (0.432)	1.743 (1.164)	- -	- -
Black	-0.310 (0.402)	-0.364 (0.299)	0.081 (0.774)	-0.050 (1.102)	0.073 (0.202)
Other race	-0.338 (0.380)	-0.499 (0.315)	0.994 (0.957)	-2.620 (7.185)	-0.158 (0.212)
Hispanic	1.399 (1.452)	1.142 (0.780)	-4.053 (3.692)	4.273 (7.512)	0.298 (0.693)
Some college	-0.045 (0.612)	0.152 (0.477)	-2.851 (1.608)	-3.656 (2.437)	0.153 (0.345)
High school	0.283 (0.469)	0.311 (0.341)	-1.190 (1.149)	-2.211 (2.403)	0.074 (0.236)
Less than high school	-0.013 (0.374)	0.049 (0.298)	-1.590 (0.972)	-1.209 (1.966)	-0.255 (0.204)
Family income	0.0003*** (0.0001)	-0.000 (0.0002)	0.0005* (0.0003)	0.0002* (0.0001)	0.0005*** (0.0002)
Household size	0.027 (0.022)	0.032 (0.021)	0.004 (0.042)	0.057 (0.040)	0.025 (0.020)
Married	0.022 (0.089)	0.060 (0.093)	0.183 (0.176)	0.027 (0.132)	0.104* (0.060)
Employed	0.040 (0.025)	0.033 (0.032)	0.008 (0.040)	0.064 (0.045)	0.012 (0.024)
Number of observations	26573	16096	10477	11451	15122
Number of instruments	42	42	42	41	41
p-value AR(1) test	0.000	0.000	0.000	0.000	0.000
p-value AR(2) test	0.509	0.063	0.498	0.897	0.356
p-value Hansen test	0.004	0.002	0.000	0.420	0.024
p-value Difference-in-Hansen test	0.146	0.001	0.242	0.058	0.932

Windmeijer finite-sample corrected standard errors are in parentheses.

All regressions include year fixed effects.

*significant at 10% level, **significant at 5% level, ***significant at 1% level

Table 6. OLS, FE and RE Estimates for Drinking Participation

Independent Variable	Full sample	Aged 50 and younger	Aged 51 and older	Males	Females
	(1)	(2)	(3)	(4)	(5)
Panel A: OLS					
Lagged drinking participation	0.629*** (0.005)	0.600*** (0.007)	0.668*** (0.008)	0.651*** (0.008)	0.611*** (0.007)
Lagged smoking participation	0.022*** (0.006)	0.024*** (0.007)	0.016 (0.010)	0.008 (0.008)	0.033*** (0.008)
Cigarette price	0.027*** (0.005)	0.024*** (0.007)	0.032*** (0.007)	0.027*** (0.007)	0.028*** (0.007)
Smoking ban index	0.0003 (0.0002)	0.0003 (0.0003)	-0.0003 (0.0003)	0.00003 (0.0003)	0.0006** (0.0003)
Beer price	-0.011*** (0.004)	-0.013** (0.005)	-0.010 (0.006)	-0.005 (0.006)	-0.016*** (0.005)
0.08 BAC	-0.031*** (0.007)	-0.028*** (0.009)	-0.038*** (0.013)	-0.021** (0.011)	-0.040*** (0.010)
Panel B: FE					
Lagged drinking participation	-0.103*** (0.009)	-0.108*** (0.012)	-0.146*** (0.016)	-0.133*** (0.015)	-0.084*** (0.012)
Lagged smoking participation	-0.011 (0.013)	-0.022 (0.016)	-0.006 (0.023)	-0.011 (0.018)	-0.007 (0.019)
Cigarette price	0.003 (0.009)	-0.017 (0.013)	0.019 (0.015)	0.002 (0.013)	0.004 (0.013)
Smoking ban index	0.0005 (0.0003)	0.0006 (0.001)	0.0001 (0.0004)	0.0003 (0.0004)	0.0004 (0.0004)
Beer price	-0.004 (0.008)	-0.012 (0.011)	0.013 (0.013)	-0.013 (0.011)	0.006 (0.012)
0.08 BAC	-0.002 (0.009)	0.003 (0.012)	-0.009 (0.016)	0.019 (0.013)	-0.020 (0.012)
Panel C: RE					
Lagged drinking participation	0.641*** [1.850] (0.005)	0.606*** [1.745] (0.007)	0.686*** [2.014] (0.008)	0.656*** [1.946] (0.008)	0.626*** [1.781] (0.007)
Lagged smoking participation	0.035*** [0.092] (0.009)	0.035*** [0.099] (0.010)	0.029* [0.073] (0.017)	0.012 [0.035] (0.012)	0.054*** [0.136] (0.013)
Cigarette price	0.044*** [0.115] (0.008)	0.035*** [0.098] (0.010)	0.056*** [0.141] (0.014)	0.042*** [0.124] (0.012)	0.045*** [0.113] (0.011)
Smoking ban index	0.0005 [0.001] (0.0003)	0.0003 [0.0009] (0.0004)	0.0006 [0.001] (0.0006)	-0.0001 [-0.0002] (0.0005)	0.0009** [0.002] (0.0005)
Beer price	-0.017*** [-0.045] (0.006)	-0.018** [-0.050] (0.008)	-0.015 [-0.039] (0.011)	-0.007 [-0.022] (0.009)	-0.024** [-0.060] (0.008)
0.08 BAC	-0.054*** [-0.143] (0.012)	-0.044*** [-0.124] (0.014)	-0.072*** [-0.183] (0.023)	-0.035** [-0.105] (0.017)	-0.067*** [-0.172] (0.017)
Number of observations	27049	16386	10663	11761	15288

Parameter estimates are in brackets and standard errors are in parentheses.

Other control variables are omitted in the table.

All regressions include year fixed effects.

*significant at 10% level, **significant at 5% level, ***significant at 1% level

Table 7. CML Estimates for Drinking Participation with Lagged Smoking Participation

Independent Variable	Full sample	Aged 50 and younger	Aged 51 and older	Males	Females
	(1)	(2)	(3)	(4)	(5)
Lagged drinking participation	0.233*** [0.634] (0.015)	0.234*** [0.676] (0.018)	0.384*** [1.006] (0.025)	0.139*** [0.486] (0.021)	0.280*** [0.719] (0.018)
Lagged smoking participation	0.005 [0.015] (0.022)	0.005 [0.015] (0.024)	0.003 [0.007] (0.040)	-0.003 [-0.012] (0.026)	0.019 [0.048] (0.031)
Cigarette price	0.042*** [0.115] (0.013)	0.021 [0.063] (0.015)	0.068*** [0.171] (0.020)	0.037** [0.138] (0.017)	0.042** [0.106] (0.018)
Smoking ban index	0.0008** [0.002] (0.0005)	0.0007 [0.002] (0.0006)	0.0007 [0.002] (0.0007)	0.0001 [0.0004] (0.0006)	0.001** [0.004] (0.0006)
Beer price	-0.020* [-0.054] (0.010)	-0.026** [-0.078] (0.011)	-0.004 [-0.011] (0.016)	-0.006 [-0.022] (0.013)	-0.029** [-0.073] (0.013)
0.08 BAC	-0.046*** [-0.128] (0.016)	-0.035** [-0.108] (0.017)	-0.079*** [-0.201] (0.030)	-0.009 [-0.033] (0.0001)	-0.074*** [-0.189] (0.023)
Age	-0.008*** (0.003)	-0.002 (0.004)	-0.021*** (0.008)	-0.010** (0.005)	-0.008 (0.005)
Male	0.096*** (0.014)	0.075*** (0.015)	0.086*** (0.021)	-	-
Black	-0.113*** (0.018)	-0.091*** (0.019)	-0.134*** (0.027)	-0.094*** (0.026)	-0.128*** (0.023)
Other race	-0.069** (0.030)	-0.045 (0.031)	-0.101** (0.051)	-0.049 (0.044)	-0.085** (0.038)
Hispanic	0.010 (0.035)	0.033 (0.033)	-0.054 (0.063)	0.017 (0.045)	0.007 (0.046)
Some college	0.032 (0.075)	0.004 (0.078)	0.205 (0.161)	0.230** (0.098)	0.006 (0.085)
High school	-0.151** (0.068)	-0.124* (0.072)	-0.177 (0.156)	-0.615 (0.470)	-0.141* (0.073)
Less than high school	-0.032 (0.067)	-0.027 (0.071)	-0.027 (0.160)	0.289*** (0.091)	-0.029 (0.075)
Family income	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Household size	0.0006 (0.008)	-0.003 (0.008)	0.010 (0.017)	-0.004 (0.010)	0.005 (0.011)
Married	-0.092*** (0.023)	-0.077*** (0.025)	-0.130*** (0.050)	-0.045 (0.030)	-0.119*** (0.033)
Employed	0.033** (0.016)	0.023 (0.018)	0.033 (0.029)	0.035* (0.020)	0.027 (0.022)
Number of observations	26950	16307	10643	11717	15233

Marginal effects are reported.

Parameter estimates are in brackets. Standard errors are in parentheses.

Other control variables are omitted in the table.

All regressions include year fixed effects.

*significant at 10% level, **significant at 5% level, ***significant at 1% level

Table 8. CML Estimates for Drinking Participation with Lagged Cigarette Consumption

Independent Variable	Full sample (1)	Aged 50 and younger (2)	Aged 51 and older (3)	Males (4)	Females (5)
Lagged drinking participation	0.228*** [0.621] (0.015)	0.225*** [0.652] (0.018)	0.387*** [1.013] (0.025)	0.134*** [0.473] (0.021)	0.274*** [0.703] (0.018)
Lagged cigarette consumption	0.001 [0.003] (0.001)	0.0005 [0.002] (0.001)	0.002 [0.004] (0.002)	0.001 [0.004] (0.001)	0.001 [0.003] (0.002)
Cigarette price	0.046*** [0.125] (0.013)	0.024 [0.073] (0.015)	0.071*** [0.180] (0.020)	0.041** [0.153] (0.017)	0.045** [0.113] (0.018)
Smoking ban index	0.0008* [0.002] (0.0005)	0.0006 [0.002] (0.0006)	0.0006 [0.002] (0.0008)	0.000 [0.0002] (0.0006)	0.001** [0.004] (0.0007)
Beer price	-0.022** [-0.060] (0.010)	-0.027** [-0.080] (0.011)	-0.008 [-0.019] (0.015)	-0.007 [-0.026] (0.012)	-0.031** [-0.079] (0.013)
0.08 BAC	-0.045*** [-0.127] (0.016)	-0.034** [-0.104] (0.017)	-0.080*** [-0.204] (0.030)	-0.007 [-0.027] (0.020)	-0.075*** [-0.191] (0.023)
Age	-0.010*** (0.003)	-0.003 (0.004)	-0.022*** (0.008)	-0.010** (0.005)	-0.011** (0.005)
Male	0.101*** (0.014)	0.080*** (0.015)	0.090*** (0.020)	-	-
Black	-0.109*** (0.018)	-0.087*** (0.019)	-0.133*** (0.027)	-0.082*** (0.025)	-0.130*** (0.023)
Other race	-0.070** (0.031)	-0.047 (0.031)	-0.096* (0.050)	-0.052 (0.044)	-0.083** (0.038)
Hispanic	0.009 (0.035)	0.032 (0.034)	-0.057 (0.062)	0.021 (0.044)	0.0009 (0.047)
Some college	0.038 (0.074)	0.005 (0.078)	0.218 (0.157)	0.227** (0.097)	0.032 (0.084)
High school	-0.146** (0.068)	-0.133* (0.071)	-0.161 (0.156)	-0.612 (0.471)	-0.137* (0.072)
Less than high school	-0.024 (0.067)	-0.021 (0.069)	-0.016 (0.160)	0.287*** (0.091)	-0.023 (0.075)
Family income	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Household size	0.0009 (0.008)	-0.002 (0.008)	0.012 (0.017)	-0.003 (0.009)	0.004 (0.011)
Married	-0.096*** (0.023)	-0.081*** (0.024)	-0.133*** (0.050)	-0.050* (0.029)	-0.123*** (0.033)
Employed	0.033** (0.016)	0.022 (0.018)	0.038 (0.028)	0.036* (0.020)	0.027 (0.022)
Number of observations	27036	16395	10641	11734	15302

Marginal effects are reported.

Parameter estimates are in brackets. Standard errors are in parentheses.

All regressions include year fixed effects.

*significant at 10% level, **significant at 5% level, ***significant at 1% level

Table 9. MSL Estimates for Drinking Participation

Panel A	Full sample	Age 50 and younger	Aged 51 and older	Males	Females
Independent Variable	(1)	(2)	(3)	(4)	(5)
Lagged drinking participation	[1.851]*** (0.019)	[1.702]*** (0.026)	[1.971]*** (0.038)	[1.971]*** (0.031)	[1.791]*** (0.025)
Lagged smoking participation	[0.092]*** (0.024)	[0.082]** (0.032)	[0.034] (0.058)	[0.038] (0.038)	[0.141]*** (0.033)
Cigarette price	[0.115]*** (0.022)	[0.107]*** (0.031)	[0.104]** (0.043)	[0.125]*** (0.036)	[0.113]*** (0.029)
Smoking ban index	[0.001] (0.0009)	[0.0008] (0.001)	[0.002] (0.002)	[-0.0003] (0.001)	[0.002]* (0.001)
Beer price	[-0.045]*** (0.017)	[-0.046]** (0.023)	[-0.054] (0.034)	[-0.026] (0.027)	[-0.059]*** (0.022)
0.08 BAC	[-0.143]*** (0.034)	[-0.102]** (0.046)	[-0.190]*** (0.066)	[-0.100]* (0.053)	[-0.169]*** (0.044)
Number of observations	27049	13957	7078	11342	15110
Panel B					
Lagged drinking participation	[1.848]*** (0.019)	[1.699]*** (0.026)	[1.974]*** (0.038)	[1.963]*** (0.031)	[1.793]*** (0.025)
Lagged cigarette consumption	[0.002] (0.001)	[0.001] (0.002)	[0.000] (0.003)	[0.001] (0.002)	[0.002] (0.002)
Cigarette price	[0.121]*** (0.022)	[0.114]*** (0.031)	[0.108]** (0.042)	[0.133]*** (0.036)	[0.115]*** (0.029)
Smoking ban index	[0.001] (0.0009)	[0.0008] (0.001)	[0.002] (0.002)	[-0.0003] (0.001)	[0.002]* (0.001)
Beer price	[-0.045]*** (0.017)	[-0.043]* (0.023)	[-0.058]* (0.034)	[-0.028] (0.027)	[-0.060]*** (0.022)
0.08 BAC	[-0.141]*** (0.033)	[-0.097]** (0.046)	[-0.192]*** (0.066)	[-0.096]* (0.053)	[-0.169]*** (0.044)
Number of observations	26959	13896	7050	11284	15076

Panel A uses lagged smoking participation as an independent variable;

Panel B uses lagged cigarette consumption.

Other control variables are omitted in the table.

All regressions include year fixed effects.

Parameter estimates are in brackets and standard errors are in parentheses.

*significant at 10% level, **significant at 5% level, ***significant at 1% level

Table 10. CML Estimates for Heavy Drinking Participation

Panel A	Full sample	Aged 50 and younger		Aged 51 and older	Males	Females
		(1)	(2)	(3)	(4)	(5)
Independent Variable						
Lagged heavy drinking participation	0.007* [0.553] (0.004)	0.018* [0.704] (0.010)	0.007 [0.636] (0.006)	0.013* [0.518] (0.007)	0.013 [0.708] (0.012)	
Lagged smoking participation	0.0008 [0.125] (0.0008)	0.002 [0.140] (0.002)	0.001 [0.247] (0.001)	0.004 [0.254] (0.003)	-0.0007 [-0.114] (0.0009)	
Cigarette price	-0.00005 [-0.078] (0.0004)	0.000 [0.006] (0.0007)	-0.001 [-0.226] (0.0006)	-0.001 [-0.103] (0.001)	-0.0003 [-0.041] (0.0006)	
Smoking ban index	0.000 [0.0001] (0.000)	-0.000 [-0.0006] (0.000)	0.000 [0.0008] (0.000)	0.000 [0.002] (0.000)	-0.000 [-0.002] (0.000)	
Beer price	-0.0003 [-0.050] (0.0003)	-0.0003 [-0.035] (0.0006)	-0.0003 [-0.056] (0.0003)	-0.0007 [-0.049] (0.0009)	-0.0004 [-0.055] (0.0005)	
0.08 BAC	-0.0001 [-0.023] (0.0005)	-0.0006 [-0.049] (0.001)	-0.0002 [-0.053] (0.0008)	-0.0004 [-0.024] (0.002)	-0.0001 [-0.022] (0.0001)	
Number of observations	26950	16307	10643	11717	15233	
Panel B						
Lagged heavy drinking participation	0.007* [0.537] (0.004)	0.020* [0.710] (0.012)	0.005 [0.563] (0.005)	0.014* [0.450] (0.007)	0.013 [0.690] (0.011)	
Lagged cigarette consumption	0.0001*** [0.018] (0.00004)	0.0002** [0.018] (0.0001)	0.000 [0.020] (0.000)	0.0003*** [0.022] (0.0001)	0.000 [0.009] (0.000)	
Cigarette price	-0.0004 [-0.060] (0.0004)	0.0002 [0.014] (0.0008)	-0.0007 [-0.202] (0.0005)	-0.001 [-0.095] (0.001)	-0.000 [-0.004] (0.001)	
Smoking ban index	0.000 [0.0003] (0.000)	0.000 [0.0002] (0.000)	-0.000 [-0.0001] (0.000)	0.000 [0.002] (0.000)	-0.000 [-0.001] (0.000)	
Beer price	-0.0003 [-0.050] (0.0003)	-0.0004 [-0.033] (0.0006)	-0.0002 [-0.637] (0.0003)	-0.0007 [-0.047] (0.0009)	-0.0004 [-0.054] (0.0005)	
0.08 BAC	-0.0001 [-0.021] (0.0006)	-0.0006 [-0.048] (0.001)	-0.0002 [-0.045] (0.0006)	-0.0003 [-0.017] (0.002)	-0.0002 [-0.028] (0.001)	
Number of observations	27036	16395	10641	11734	15302	

Panel A uses lagged smoking participation as an independent variable;

Panel B uses lagged cigarette consumption.

Other control variables are omitted in the table.

All regressions include year fixed effects.

Marginal effects are reported. Standard errors are in parentheses.

*significant at 10% level, **significant at 5% level, ***significant at 1% level

Table 11. MSL Estimates for Heavy Drinking Participation

Panel A Independent Variable	Full sample (1)	Age 50 and younger	Aged 51 and older	Males	Females
		(2)	(3)	(4)	(5)
Lagged heavy drinking participation	[1.766]*** (0.048)	[1.589]*** (0.070)	[2.099]*** (0.096)	[1.769]*** (0.057)	[1.805]*** (0.096)
Lagged smoking participation	[0.421]*** (0.038)	[0.430]*** (0.052)	[0.354]*** (0.093)	[0.435]*** (0.048)	[0.413]*** (0.067)
Cigarette price	[-0.022] (0.042)	[-0.005] (0.058)	[-0.144] (0.090)	[-0.064] (0.053)	[0.034] (0.069)
Smoking ban index	[-0.0004] (0.002)	[-0.0001] (0.002)	[0.006] (0.004)	[-0.0009] (0.002)	[-0.0009] (0.003)
Beer price	[-0.027] (0.032)	[-0.022] (0.044)	[-0.030] (0.067)	[-0.060] (0.040)	[-0.023] (0.054)
0.08 BAC	[-0.059] (0.061)	[-0.066] (0.086)	[-0.100] (0.122)	[-0.069] (0.075)	[-0.021] (0.108)
Number of observations	27049	13957	7078	11342	15110
Panel B					
Lagged heavy drinking participation	[1.749]*** (0.048)	[1.563]*** (0.071)	[2.086]*** (0.097)	[1.743]*** (0.057)	[1.809]*** (0.096)
Lagged cigarette consumption	[0.019]*** (0.002)	[0.022]*** (0.003)	[0.015]*** (0.004)	[0.019]*** (0.002)	[0.020]*** (0.004)
Cigarette price	[0.012] (0.042)	[-0.001] (0.058)	[-0.132] (0.091)	[-0.053] (0.054)	[0.038] (0.068)
Smoking ban index	[-0.0003] (0.002)	[0.0007] (0.002)	[0.005] (0.004)	[-0.001] (0.002)	[-0.0005] (0.003)
Beer price	[-0.027] (0.031)	[-0.018] (0.044)	[-0.041] (0.068)	[-0.058] (0.040)	[-0.026] (0.054)
0.08 BAC	[-0.057] (0.061)	[-0.064] (0.085)	[-0.099] (0.122)	[-0.061] (0.075)	[-0.031] (0.108)
Number of observations	27058	13951	7069	11326	15132

Panel A uses lagged smoking participation as an independent variable;
Panel B uses lagged cigarette consumption.
Other control variables are omitted in the table.
All regressions include year fixed effects.
Parameter estimates are in brackets and standard errors are in parentheses.
*significant at 10% level, **significant at 5% level, ***significant at 1% level

References

- Abel, E L, Kruger, M L & Friedl, J. "How do physicians define "light," "moderate," and "heavy" drinking?" *Alcoholism, clinical and experimental research* 1998,22(5):979-84.
- Arellano, M., & Bond S. "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." *The Review of Economic Studies*, April 1991, 58, pp. 277 – 297.
- Arellano, M., & Bover, O. "Another look at the instrumental-variable estimation of error- components models." *Journal of Econometrics*, 1995, 68, 29-52.
- Bask, M., & Melkersson, M. "Rationally addicted to drinking and smoking?" *Applied Economics*, 2004, 36(4), 373–381.
- Becker, G., Grossman, M., & Murphy K. "An empirical analysis of cigarette addiction." *The American Economic Review*, 1994, 84(3): 396–418.
- Becker, G., & Murphy, K. "A theory of rational addiction." *The Journal of Political Economy*, 1998, 96(4): 675-700.
- Blundell, R., & Bond, S. "Initial conditions and moment restrictions in dynamic panel data models." *Journal of Econometrics*, 1998, 87: 115–143.
- Bond, S. "Dynamic panel data models: a guide to micro data methods and practice." Institute for Fiscal Studies, WP09/02, 2002.
- Bond, S., & Windmeijer, F. "Finite sample inference for GMM estimators in linear panel data models." Institute of Fiscal Studies, WP02/x, 2002.
- Cameron, C., & Trivedi, P. *Microeconometrics: Methods and Applications*. New York, NY: Cambridge University Press, 2005
- Cameron, L., & Williams, J. "Cannabis, alcohol and cigarettes: Substitutes or complements?" *The Economic Record*, 2001, 77(236), 19–34.
- Chaloupka, F.J., & Grossman, M. "Price, tobacco control policies and youth smoking." National Bureau of Economic Research, Inc, NBER Working Papers 5740, 1996.
- Chaloupka, F.J., & Warner, K.E. "The Economics of Smoking." In *The Handbook of Health Economics*, ed. Joseph P. Newhouse and Anthony J. Cuyler, 2000.
- Chamberlain, G. "Panel data." In Handbook of Econometrics, Griliches Z, Intriligator M (eds), vol. 2. North Holland: Amsterdam, 1984, 1247–1318.
- Chay, Kenneth Y. and Dean Hyslop (1998), "Identification and Estimation of Dynamic Binary Response Panel Data Models: Empirical Evidence using Alternative Approaches." Center for Labor Economics, Working Paper No. 5, UC Berkeley.
- Contoyannis, P., Jones, A., & Rice, N. "Simulation-based inference in dynamic panel probit models: An application to health." *Empirical Economics*, 2004, Vol. 29 Issue 1, pp. 49-77.

- Dave, D. M., & Saffer, H. "Alcohol Demand and Risk Preference." *Journal of Economic Psychology*, 2008, 29 (6), 810-831.
- Dee, T. "The complementarity of teen smoking and drinking." *Journal of Health Economics*, 1999, 18(6), 769-793.
- Decker, S., & Schwartz, A. "Cigarettes and alcohol: Substitutes or complements?" National Bureau of Economic Research, Inc, NBER Working Papers 7535, 2000.
- Dorn, M., et al. "Alcohol drinking patterns and non-fatal myocardial infarction in women." *Addiction*, 2007, 102(5), 730-739.
- Farchi, G., et al. "Alcohol and survival in the Italian rural cohorts of the Seven Countries Study." *International Journal of Epidemiology*, 2000, 29, 667-671.
- Gaziano, J., et al., "Light-to-moderate alcohol consumption and mortality in the Physicians' Health Study enrollment cohort." *Journal of the American College of Cardiology*, 2000, 35(1), 96-105.
- Geweke, J. "Bayesian inference in econometric models using Monte Carlo integration." *Econometrica*, 1989, 57(6), 1317-1339.
- Goel, K., & Morey, M. "The interdependence of cigarette and liquor demand." *Southern Economic Journal*, October 1995, Vol. 62, No. 2, pp. 451-459.
- Greene, W. *Econometric Analysis* (5th ed). Upper Saddle River, NJ: Pearson Education, Inc. 2003.
- Grossman, M., et al. "Policy watch: Alcohol and cigarette taxes." *The Journal of Economic Perspective*, Autumn 1993, Vol. 7, No. 4, pp. 211-222.
- Hajivassiliou, V. "Smooth simulation estimation of panel data LDV models." Unpublished Manuscript, Yale University, 1990.
- Heckman, J. 1981. "The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process." In *Structural Analysis of Discrete Data and Econometric Applications*, ed. C. F. Manski and D. L. McFadden, 178-195. Cambridge, MA: MIT Press.
- Hayashi, F. *Econometrics*. Princeton, NJ: Princeton University Press, 2000.
- Holtz-Eakin, D., Newey, W., & ROSEN, H. "Estimating Vector Autoregressions with Panel Data." *Econometrica*, 1988, 56, 1371-1395.
- Hsiao, C. *Analysis of Panel Data*. New York, NY: Cambridge University Press, 2003.
- Keane M. "A computationally practical simulation estimator for panel data." *Econometrica*, 1994, 62, 95-116.
- Kenkel, D., Mathios, A., & Pacula, R. "Economics of youth drug use, addiction and gateway effects." *Addiction*, 2001, 96, 151-164.

- Manson, J. E., et al. *Prevention of Myocardial Infarction*. New York: Oxford University Press, 1996.
- Maraldi, C., et al. "Impact of inflammation on the relationship among alcohol consumption, mortality, and cardiac events: the Health, Aging, and Body Composition Study." *Archives of Internal Medicine*, 2006, 166(14), 1490-1497.
- Markowitz, S., & Tauras, J. "Substance use among adolescent students with consideration of budget constraints." *Review of Economics of the Household*, 2009, 7:423-446.
- Maskarinec, G., et al. "Alcohol intake, body weight, and mortality in a multiethnic prospective cohort." *Epidemiology*, 1998, 9(6), 654-661.
- Mundlak, Y. "On the pooling of time series and cross section data." *Econometrica*, 1978, 46: 69-85.
- National Institute on Alcohol Abuse and Alcoholism. *Alcohol Alert*, No. 45. October, 1999.
- Nickell, S. "Biases in dynamic models with fixed effects." *Econometrica*, 1981, 49(6): 1417-26.
- Orzechowski, W. & Walker R. *The tax burden on tobacco: Historical compilation*. Arlington, Virginia, 2008.
- Pacula, R. "Economic modeling of the gateway effect." *Health Economics*, 1997, 6: 521-524.
- Picone, G., Sloan, F., & Trogdon, J. G. "The effect of the tobacco settlement and smoking bans on alcohol consumption." *Health Economics*, 2004, 13(10), 1063-1080.
- Rimm, B., et al. "Prospective study of alcohol consumption and risk of coronary disease in Men." *The Lancet*, 1991, 338, 464-68.
- Roodman, D. "How to do xtabond2: an introduction to "Difference" and "System" GMM in Stata." Center for Global Development Working Paper Number 103, December 2006.
- Sesso, H.D., et al. "Seven-year changes in alcohol consumption and subsequent risk of cardiovascular disease in men." *Archives of Internal Medicine*, 2001, 160, 2505-2612.
- State University of New York University at Buffalo. Moderate Drinking Lowers Women's Risk of Heart Attack. Press release, May 23, 2007.
- Simons, L., et al. "Moderate alcohol intake is associated with survival in the elderly: the Dubbo Study." *The Medical Journal of Australia*, 2000, 172, 121-124
- Stewart, M. "Maximum Simulated Likelihood Estimation of Random Effects Dynamic Probit Models with Autocorrelated Errors." *Stata Journal*, 2006, 6(2), 256-272.
- Stewart, M. "The Inter-related Dynamics of Unemployment and Low-Wage Employment." *Journal of Applied Econometrics*, April 2007, Vol.22, No. 3, pp. 511-531.
- Train, K. *Discrete Choice Methods with Simulation*. Cambridge: Cambridge University Press, 2003.

Windmeijer, F. "A finite sample correction for the variance of linear two-step GMM Estimators." *Journal of Econometrics*, 2005, 126: 25-51.

Wooldridge, J. *Econometric Analysis of Cross Section and Panel Data*. MIT Press: Cambridge, MA, 2002.

Wooldridge, J. "Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity." *Journal of Applied Econometrics*, 2005, 20: 39-54.

Zhao, X., & Harris, M. "Demand for marijuana, alcohol and tobacco: Participation, levels of consumption and cross-equation correlations." *The Economic Record*, 2004, 80(251), 394-410.