Addiction and Social Interactions: Theory and Evidence

Julian Reif*

University of Chicago

February 14, 2012

Abstract

Many activities such as smoking and gambling have long been subject to government regulation. Congress has also considered passing laws to encourage healthy eating and some corporations subsidize exercising. Each of these activities can be addictive and is more enjoyable when done with others. These features, addiction and social interactions, make regulation potentially important but difficult to design. This paper develops and estimates the first model of demand for a good that is both addictive and affected by social interactions. A key benefit of jointly modeling these two phenomena is that addiction introduces dynamics that allow the researcher to identify social interactions under relatively weak assumptions. The model also illustrates that the type of externalities that arise depends greatly on the precise manner in which social interactions affect individual utility. Finally, I apply my proposed identification methodology to data on cigarette consumption from the National Health Interview Survey and find strong evidence for the presence of both addiction and social interactions. Accounting for social interactions together with addiction triples the estimated effect of anti-smoking laws on consumption compared to a model that accounts for neither.

*I thank Steven Durlauf, Ali Hortacsu, Anup Malani, and Derek Neal for continual support and helpful comments. I also thank Jason Abaluck, Gary Becker, Tatyana Deryugina, Hays Golden, Gabriel Ulyssea, and participants at the University of Chicago Labor Working Group and the University of Chicago Microeconomics Lunch for useful discussions and feedback. I am grateful to the National Science Foundation for financial support.

1 Introduction

Governments have long regulated activities believed to have harmful consequences for consumers. For example, the U.S. state and federal governments discourage cigarette smoking by levying large excise taxes, many states ban commercial casinos, and Congress has considered implementing a "fat tax" to reduce obesity. Corporations, too, are interested in promoting healthy behavior: IBM employees receive a $150 bonus for exercising (The Economist 2011).

Each of these activities is affected by social interactions. More precisely, the utility an individual receives from consuming these activities depends directly on the consumption of other individuals. A large theoretical and empirical literature has examined social interactions in a wide variety of contexts.[1] Many of these activities can also be addictive: current consumption increases the utility of future consumption. The presence of addiction in consumer behavior has led to a separate literature on modeling the addiction process.[2]

This paper develops and estimates the first model of demand for a good that is both addictive and affected by social interactions. A key benefit of jointly modeling these two phenomena is that addiction introduces dynamics that allow the researcher to identify social interactions, which are often difficult or impossible to identify in static models. I also demonstrate that identifying the presence of social interactions is not sufficient to infer their form, for different types of social interactions can generate indistinguishable demand equations. Finally, I apply my proposed identification methodology to data on cigarette consumption and find strong evidence for the presence of both addiction and social interactions.

I explicitly model the individual decision problem in order to illustrate the precise effect different forms of social interactions have on individual behavior. I consider two different forms of interactions: "conformity"and "proportional spillovers". Under conformity, individuals face incentives to match their consumption to the mean of

[1]A less than comprehensive list includes the following. For theory, see Bernheim (1994), Binder and Pesaran (2001), Bisin, Horst, and Ozgur (2006), Blume, Brock, Durlauf, and Ioannides (2010), Brock and Durlauf (2001), and Manski (1993). Empirical analyses of social interactions have examined their effect on crime (Glaeser, Sacerdote, and Scheinkman 1996), disadvantaged youth (Case and Katz 1991), grades (Sacerdote 2001), obesity (Christakis and Fowler 2007), and smoking (Gaviria and Raphael 2001; Krauth 2007; Nakajima 2007; Powell, Tauras, and Ross 2005).

[2]See Becker, Grossman, and Murphy (1994), Becker and Murphy (1988), Chaloupka (1991), Gruber and Koszegi (2001), Levy (2010), O'Donoghue and Rabin (2002), and Orphanides and Zervos (1995).

their reference group. Under proportional spillovers, individual utility is affected linearly by a change in the mean consumption level.

The model reveals that the effect of social interactions on consumption depends greatly on their form. An increase in conformity causes individuals to place more weight on the average consumption in their group and less weight on their own idiosyncratic preferences. This compresses the distribution of consumption within the group: higher conformity causes individuals with preferences for a high level of consumption to consume less and causes individuals with preferences for a low level of consumption to consume more. There is no effect on average consumption because the effect on high-consuming individuals is canceled out by an opposite effect on low-consuming individuals. Proportional spillovers, by contrast, increase everyone's consumption by a constant factor.

Despite their disparate effects on consumption, both conformity and proportional spillovers imply a linear relationship between individual and mean group consumption when utility is quadratic. Thus both these forms are consistent with the "linear-in-means" model estimated by many social interactions studies.

It is well known that it is difficult to identify social interactions in the linear-in-means model (Manski 1993). Consistent estimation via ordinary least squares requires the absence of unobserved group-wide shocks, which is unlikely to hold in most empirical analyses. This paper shows how the dynamics that arise naturally as a byproduct of coupling social interactions with addiction allow researchers to identify social interactions even in the presence of unobserved group-wide shocks. The intuition follows from two simple observations. First, addiction causes group consumption to be related to past determinants of demand. Second, although shocks to demand are likely to be correlated across different individuals within the same reference group, it is unlikely that those shocks are correlated with past determinants of demand. In other words, past exogenous determinants of demand are valid instruments for identifying social interactions. If individuals are forward looking then future determinants of demand are also valid instruments.

Finally, I apply my model to data on cigarette consumption–which medical and economic evidence suggest is both addictive and affected by social interactions–obtained from the National Health Interview Survey. I define a smoker's reference group to be other smokers who live in her metropolitan area and are from the same demographic group.[3] I exploit a key implication of the model–that current consumption is related to past determinants of demand–to identify the relationship between group and individual cigarette consumption. My main identification assumption is that lagged cigarette prices and lagged anti-smoking laws are uncorrelated with current unobserved determinants of smoking demand. I show that my results are robust to the possibility that *current* cigarette prices are correlated with unobserved determinants of smoking demand.

My results demonstrate a strong and statistically significant relationship between individual and group consumption. The interpretation of these results depends greatly on the assumed form of social interactions. Assuming conformity, these results imply that social interactions significantly reduce the variance of individual consumption.[4] If one assumes that proportional spillovers rather than conformity is the proper specification, however, then these results imply that an increase of one cigarette per day in a group's consumption of cigarettes is associated with an increase in consumption of approximately 0.25 cigarettes per day for a current smoker.

This paper makes three contributions to the literature. First, it derives a novel model of demand that accounts for both the dynamic aspects of addiction and the endogenous effects of interacting consumers. The model shows that policy makers need to identify not just the presence but also the precise form of social interactions in order to infer the type of externalities generated by those interactions. For example, conformity has no effect on the average level of consumption while proportional spillovers increase the level by a constant factor.

Second, this paper is the first to show how the dynamics inherent in such a model can be used to identify the presence of endogenous social interactions. There is a large literature concerned with how to identify social interactions.[5] It is usually difficult to find an appropriate instrument in static settings: variables that affect group consumption generally must also affect individual consumption and thus cannot be used as excluded instruments. My proposed method, by contrast, has sound theoretical justification. Lagged exogenous determinants of demand are necessarily related

[3] I allow the demographic group to vary alternately by age, gender, and marital status.

[4] Quantifying the size of the reduction requires information about the distribution of individual determinants of demand. Reasonable assumptions imply a reduction of approximately forty percent in the standard deviation of consumption.

[5] See Blume, Brock, Durlauf, and Ioannides (2010), Durlauf and Tanaka (2008), Graham and Hahn (2005), and Manski (1993).

to group consumption in the model because of addiction. Yet, they are uncorrelated with the unobserved determinants of individual demand if the researcher has controlled for the individual's stock of past consumption because lagged determinants of demand affect the individual only through that stock.

Third, I present empirical evidence that an individual's cigarette consumption is strongly affected by the average consumption in her reference group. This paper is the first to estimate this relationship.[6] I also find that the recent decline in smoking consumption is strongly associated with anti-smoking laws as measured by a comprehensive anti-smoking laws index. Moreover, a naive model underestimates the elasticity of demand with respect to this index by a factor of three as compared to a model that properly accounts for addiction and social interactions.

The rest of this paper is organized as follows. Section 2 presents the model and main theoretical results. I discuss identification in Section 3 and describe the data in Section 4. Section 5 describes the empirical strategy and Section 6 presents results. Section 7 concludes.

2 Model

A consumer's instantaneous utility is represented by the function

$$V\left(a_{it}, S_{it}, x_{it}, c_{it}, E_t\left[\overline{a}_t\right]\right)$$

where a_{it} is consumer i's consumption of an addictive good in period t, S_{it} is the consumer's stock of past consumption, x_{it} represents other components that affect utility (e.g., education or advertising), and c_{it} represents a composite good. $E_t\left[\overline{a}_t\right]$ is the consumer's time-t expectation of the mean consumption of the addictive good by other consumers in her reference group. I assume throughout that the reference group is "large" in the sense that an individual's contribution to the mean is negligible. The stock of past consumption evolves according to the equation

$$S_{it+1} = (1-d)\left(S_{it} + a_{it}\right) \tag{1}$$

[6] Most empirical papers on social interactions and smoking examine the effect of group *prevalence* on the probability an individual smokes rather than the effect of group *consumption* on individual consumption. Aristei and Pieroni (2009) and Jones (1989) estimate the effect of group smoking prevalence, but not consumption, on individual consumption.

where $d \in (0,1)$ is the rate of depreciation. Many addictive goods such as cigarettes have a negative effect on the consumer's health. I model this by assuming that, all things equal, a higher stock of past consumption lowers utility, i.e., $\partial V/\partial S_{it} < 0$.[7]

Two definitions will be helpful when discussing this model. A consumer is *addicted* if, all things equal, an increase in the consumer's stock of past consumption (S_{it}) increases her equilibrium consumption (a_{it}). The magnitude of this increase measures the strength of addiction. *Adjacent complementarity* is the effect of last period's consumption (a_{it-1}) on current equilibrium consumption (a_{it}). See Becker and Murphy (1988) for a discussion of the relationship between addiction and adjacent complementarity and of how they relate to other medical and psychological concepts of addiction.

The consumer's problem is

$$\max_{a_{it},c_{it}} \sum_{t=1}^{\infty} \beta^{t-1} V\left(a_{it}, S_{it}, x_{it}, c_{it}, E_t\left[\overline{a}_t\right]\right) \tag{2}$$

where S_{i1} is given and $\beta < 1$ is the consumer's discount rate. The consumer's budget constraint is

$$A_{i0} = \sum_{t=1}^{\infty} (1+r)^{-(t-1)} \left(c_{it} + p_t a_{it}\right) \tag{3}$$

where A_{i0} is the present value of wealth, r is the interest rate, and p_t denotes the price of the addictive good. The composite good c_{it} is taken as numeraire. I assume the consumer's discount rate is equal to $1/(1+r)$. All variables in this problem vary at the individual level except for price, mean consumption, and the discount and interest rates. I drop the i subscript for the remainder of this section for notational ease.

I impose the following assumptions in order to render this model analytically tractable:

A1 Instantaneous utility can be decomposed into a private component and a social component:

$$V\left(a_t, S_t, x_t, c_t, E_t\left[\overline{a}_t\right]\right) = U\left(a_t, S_t, x_t, c_t\right) + G\left(a_t, E_t\left[\overline{a}_t\right]\right) \tag{4}$$

[7] One could alternatively consider cases where the stock of past consumption raises utility, e.g., a consumer who is addicted to exercising. This would not affect the form of the demand equation although it could affect the signs of the regressors. Most applications of addiction, however, including the one in this paper (cigarette consumption), are concerned with cases where the stock lowers utility.

A2 Private utility is concave and quadratic:

$$U(a_t, S_t, x_t, c_t) = -\frac{1}{2}\left(u_{aa}a_t^2 + u_{ss}S_t^2 + u_{xx}x_t^2 + u_{cc}c_t^2\right) + u_{as}a_tS_t \tag{5}$$
$$+ u_{ax}a_tx_t + u_{ac}a_tc_t + u_{sx}S_tx_t + u_{sc}S_tc_t + u_{xc}x_tc_t$$
$$+ u_aa_t + u_sS_t + u_xx_t + u_cc_t$$

This functional form captures the dynamic aspects of the model and delivers linear first order conditions that aid analysis. The coefficients u_{aa}, u_{ss}, u_{xx} and u_{cc} are positive due to the assumed concavity of U. The coefficient u_{as} is assumed to be positive so that the marginal utility of consumption of a_t is higher when the addictive stock S_t is higher.

A3 Consumers have perfect foresight.

Assumption A1 is common in the social interactions literature while assumptions A2 and A3 are standard in the literature on addiction.[8] Some critics argue that perfect foresight implies that addicts should never express regret and that there is no scope for public policy (Akerlof 1991; Winston 1980). Gruber and Koszegi (2001) show, however, that perfect foresight is compatible with time inconsistency. Furthermore, Orphanides and Zervos (1995) relax the assumption of perfect foresight and show that the main empirical predictions of the Becker-Murphy addiction model remain unchanged.

My theoretical model is closely related to Bisin, Horst, and Ozgur (2006), who analyze the existence and form of equilibria in dynamic linear economies that exhibit both habit formation and social interactions. My analysis differs from theirs in the following ways. First, they assume habit formation is captured completely by last period's decision, which is equivalent to assuming the stock of past consumption takes the special form $S_{t+1} = a_t$.[9] Second, I derive equilibria in the context of a standard consumer's problem and thereby relate my analysis directly to the addiction literature.

[8] Blume, Brock, Durlauf, and Ioannides (2010) and Brock and Durlauf (2001) are good examples from the social interactions literature. Becker and Murphy (1988), Becker, Grossman, and Murphy (1994), Chaloupka (1991) and Gruber and Koszegi (2001) are prominent examples from the addiction literature.

[9] Becker and Murphy (1988) point out that in the special case where $S_{t+1} = a_t$, a positive effect of past consumption on the marginal utility of current consumption is both necessary and sufficient to induce addiction. This is no longer true in the more general case where $S_{t+1} = (1-d)(S_t + a_t)$, which is what I consider in my model.

This allows me, for example, to calculate price elasticities, unlike Bisin, Horst, and Ozgur (2006), who do not include a budget constraint. Finally, I consider two different parametric specifications for social interactions while Bisin, Horst, and Ozgur (2006) only consider one.

Solving the first order condition for c_t and substituting the result into (5) allows one to rewrite the consumer's problem (2) as a maximization problem in a_t only:

$$\max_{a_t} \sum_{t=1}^{\infty} \beta^{t-1} V^* \left(a_t, S_t, x_t, E_t\left[\overline{a}_t\right]\right) \tag{6}$$

where

$$\begin{aligned} V^* \left(a_t, S_t, x_t, E_t\left[\overline{a}_t\right]\right) = & -\frac{1}{2}\left(b_{aa}a_t^2 + b_{ss}S_t^2 + b_{xx}x_t^2\right) \\ & + b_{as}a_tS_t + b_{ax}a_tx_t + b_{sx}S_tx_t \\ & b_aa_t + b_sS_t + b_xx_t + G\left(a_t, E_t\left[\overline{a}_t\right]\right) \end{aligned} \tag{7}$$

The coefficients in (7) capture the effect of the input variables on the consumer's utility assuming optimal consumption of the composite good c_t. I show in the appendix that the coefficients on the squared terms in (7) are positive. The coefficient b_{as} measures the effect of the stock of past consumption on the marginal utility of current consumption assuming optimal consumption of c_t and is assumed to be positive.[10]

Following Brock and Durlauf (2001), I consider two different parametric representations of social utility. The first embodies a pure conformity effect:

$$G\left(a_t, E_t\left[\overline{a}_t\right]\right) = -\frac{1}{2}b_g\left(a_t - E_t\left[\overline{a}_t\right]\right)^2 \tag{8}$$

The second is a proportional spillovers effect, defined as a linear interaction between individual consumption and expected average consumption:

$$G\left(a_t, E_t\left[\overline{a}_t\right]\right) = b_ga_tE_t\left[\overline{a}_t\right] \tag{9}$$

The strength of social interactions is captured in both cases by the parameter $b_g > 0$.

[10]The appendix shows that assuming $u_{ac}u_{sc} \geq 0$ is sufficient (but not necessary) to ensure that $b_{as} > 0$. In words, $u_{ac}u_{sc} \geq 0$ means the addictive good a_t and the addictive stock S_t do not have opposing effects on the marginal utility of the composite good c_t.

Expanding (8) reveals the relationship between these two forms of social utility:

$$-\frac{1}{2}b_g\left(a_t - E_t\left[\overline{a}_t\right]\right)^2 = b_g a_t E_t\left[\overline{a}_t\right] - \frac{1}{2}b_g\left(\left(a_t\right)^2 + \left(E_t\left[\overline{a}_t\right]\right)^2\right) \tag{10}$$

Brock and Durlauf (2001) show that (8) and (9) result in the same choice problem for the individual when a_t is binary.[11] This is not so in this model because a_t is continuous.

Intuitively, an increase in conformity causes individuals to place more weight on the average consumption in their group and less weight on their own idiosyncratic preferences. This compresses the distribution of consumption within the group: higher conformity causes individuals with preferences for a high level of consumption to consume less and causes individuals with preferences for a low level of consumption to consume more. There is no effect on average consumption because the effect on high-consuming individuals is canceled out by an opposite effect on low-consuming individuals. Proportional spillovers, by contrast, increase everyone's utility (and consumption) in proportion to her current consumption. The next two subsections demonstrate these results formally by deriving demand equations for the cases where social utility takes the forms (8) and (9), respectively.

2.1 Conformity

Let social interactions be captured by a pure conformity effect:

$$G\left(a_t, E_t\left[\overline{a}_t\right]\right) = -\frac{1}{2}b_g\left(a_t - E_t\left[\overline{a}_t\right]\right)^2$$

I first consider a baseline case where consumers are myopic. Analytically, this corresponds to ignoring any forward-looking terms in the first order condition when solving the consumer's problem (6) subject to the law of motion for the stock of past consumption (1) and the budget constraint (3). This yields the following demand equation:

$$a_t = \alpha_m S_t + \gamma_m \overline{a}_t + \pi_m p_t + \delta_m x_t + k_m \tag{11}$$

[11] To see this, let $a_t \in \{-1, 1\}$. Then the right-hand side of equation (10) simplifies to $b_g a_t E_t\left[\overline{a}_t\right] - \frac{1}{2}b_g\left(1 + \left(E_t\left[\overline{a}_t\right]\right)^2\right)$. This differs from the right-hand side of (9) only by the tem $-\frac{1}{2}b_g\left(1 + \left(E_t\left[\overline{a}_t\right]\right)^2\right)$, which drops out of the consumer's maximization problem.

where

$$\alpha_m = \frac{b_{as}}{b_g + b_{aa}} > 0$$

$$\gamma_m = \frac{b_g}{b_g + b_{aa}} > 0$$

$$\pi_m = \frac{-\lambda}{b_g + b_{aa}} < 0$$

$$\delta_m = \frac{b_{ax}}{b_g + b_{aa}} \lesseqgtr 0$$

$$k_m = \frac{b_a}{b_g + b_{aa}} > 0$$

and λ is the marginal utility of wealth. Myopic consumers are necessarily addicted because $\alpha_m > 0$. Individual consumption is positively related to group consumption ($\gamma_m > 0$) and negatively related to price ($\pi_m < 0$). It is not affected by the negative consequences of addiction (as captured by b_s and b_{ss}) or by any future determinants of demand. An increase in the strength of conformity has a negative effect on the magnitude of all parameters except γ_m.

Relaxing the assumption that consumers are myopic yields the forward-looking demand equation:

$$a_t = \alpha_1 S_t + \alpha_2 a_{t+1} + \gamma_1 \overline{a}_t + \gamma_2 \overline{a}_{t+1} + \pi_1 p_t + \pi_2 p_{t+1} + \delta_1 x_t + \delta_2 x_{t+1} + k \tag{12}$$

where

$$
\begin{aligned}
\alpha_1 &= \frac{b_{as} - (1-d)^2 \beta (b_{as} + b_{ss})}{\Delta} \lesseqqgtr 0 \\
\alpha_2 &= (1-d)\beta \frac{b_g + b_{aa} + b_{as}}{\Delta} > 0 \\
\gamma_1 &= \frac{b_g}{\Delta} > 0 \\
\gamma_2 &= -(1-d)\beta \frac{b_g}{\Delta} < 0 \\
\pi_1 &= -\frac{\lambda}{\Delta} < 0 \\
\pi_2 &= (1-d)\beta \frac{\lambda}{\Delta} > 0 \\
\delta_1 &= \frac{b_{ax}}{\Delta} \lesseqqgtr 0 \\
\delta_2 &= (1-d)\beta \frac{b_{sx} - b_{ax}}{\Delta} \lesseqqgtr 0 \\
k &= \frac{b_a - (1-d)\beta (b_a - b_s)}{\Delta} \lesseqqgtr 0 \\
\Delta &= b_g + b_{aa} + (1-d)^2 \beta (b_{as} + b_{ss}) > 0
\end{aligned}
$$

See the appendix for a full derivation.

Unlike the case where consumers are myopic, the relationship between present consumption and the stock of past consumption (α_1) is ambiguous when consumers are forward looking. This ambiguity arises because forward-looking consumers account for the future harmful effects of consumption as captured by the utility parameter b_{ss}. Another difference is that current consumption is affected by future consumption when consumers are forward looking. The coefficient on future consumption measures the strength of adjacent complementarity.

The model predicts that consumption in demand equation (12) is negatively related to the current price but positively related to future price. This positive relationship may seem surprising since past and future consumption are complementary with present consumption. However, as explained in Becker, Grossman, and Murphy (1990), the demand equation (12) holds future consumption constant, eliminating the mechanism through which past and future prices affect present consumption. For example, if the future price increases but future consumption is unchanged, then some other force must be offsetting the price effect by raising the future stock. This in turn means current consumption must be higher because the future stock is directly related to both the current stock and current consumption.

The coefficient on current group consumption (γ_1) is positive in (12). This implies that current individual and current group consumption are positively related, as expected. Current individual consumption is negatively related to future group consumption for the same reason it is positively related to future price, which was discussed above. Note that all coefficients in (12) are functions of the strength of conformity (b_g). For example, the magnitude of the strength of addiction (α_1) is negatively affected by an increase in conformity.

The strength of adjacent complementarity (α_2) depends on how the stock of past consumption affects the marginal utility of current consumption (b_{as}), on how quickly a rise in addictive stock causes harm (b_{ss}), on how quickly the marginal utility of consumption diminishes (b_{aa}), and on the strength of conformity (b_g):

$$\frac{\partial \alpha_2}{\partial b_{as}} = (1-d)\,\beta \frac{b_{aa} + b_g - (b_{aa} + b_g - b_{ss})\,(1-d)^2\,\beta}{\Delta^2} > 0 \tag{13}$$

$$\frac{\partial \alpha_2}{\partial b_{ss}} = -\,(1-d)^3\,\beta^2 \frac{b_{aa} + b_{as} + b_g}{\Delta^2} < 0 \tag{14}$$

$$\frac{\partial \alpha_2}{\partial b_g} = \frac{\partial \alpha_2}{\partial b_{aa}} = (1-d)\,\beta \frac{(1-d)^2\,\beta\,(b_{as} + b_{ss}) - b_{as}}{\Delta^2} \lesseqgtr 0 \tag{15}$$

The model retains the general results from the addiction literature and also generates some new ones. Equations (13) and (14) illustrate the well-known results from Becker and Murphy (1988): adjacent complementarity increases the more past consumption raises the marginal utility of current consumption and decreases the more quickly the harm from past consumption increases. Forward-looking individuals temper their consumption of addictive goods because they anticipate the negative future consequences of a high stock of past consumption.

Equation (15) reveals that the effect of conformity on adjacent complementarity is ambiguous and that it depends on the relative magnitudes of b_{as} and b_{ss}. On the one hand, conformity has a negative effect on adjacent complementarity if prior consumption has a strong effect on current consumption (b_{as} is large). On the other hand, conformity has a positive effect on adjacent complementarity if the accumulated harm from past consumption increases quickly (b_{ss} is large).

Equation (12) is a second-order linear difference equation in S_t. The solution reveals precisely how individual consumption is affected by both past and future exogenous determinants of demand and past and future group consumption:

$$S_t = K_1 \sum_{j=1}^{\infty} (\phi_1)^{-j} h_{t+j} + K_2 \sum_{j=0}^{t-1} (\phi_2)^j h_{t-j} + (\phi_2)^t \left(S_0 - K_1 \sum_{j=1}^{\infty} (\phi_1)^{-j} h_j \right) \tag{16}$$

where $h_t = \gamma_1 \overline{a}_t + \gamma_2 \overline{a}_{t+1} + \pi_1 p_t + \pi_2 p_{t+1} + \delta_1 x_t + \delta_2 x_{t+1} + k$; S_0 is the initial stock of past consumption; K_1 and K_2 are positive constants; and ϕ_1 and ϕ_2 are the (positive) inverses of the characteristic roots of the difference equation.[12] The first term in equation (16) is a weighted average of future determinants of demand (p_{t+j} and x_{t+j}) and future group consumption. The second term is a weighted average of past determinants of demand and past group consumption. The third term represents the effect of an individual's initial condition and fades to zero over time. The parameters ϕ_1 and ϕ_2 dictate the sign and magnitude of the effect of a shock to past or future consumption on current consumption. These shocks can be changes to any factor affecting the demand for the addictive good, e.g., prices.

To see the effect of conformity on the average long-run equilibrium level of consumption, consider the simple case where there is no variance in p_t or x_t (so $p_t = p \forall t$ and $x_t = x \forall t$). The long-run equilibrium stock of past consumption is then equal to $S^* = (1-d)(S^* + a^*)$ for all consumers. This allows one to simplify (12) to

$$a^* = \pi^* p + \delta^* x + k^* \tag{17}$$

where

$$\pi^* = -\frac{d(1-\beta+d\beta)\lambda}{\xi}$$
$$\delta^* = \frac{db_{ax} + d(1-d)\beta(b_{sx} - b_{ax})}{\xi}$$
$$k^* = \frac{d(1-(1-d)\beta)b_a + d(1-d)\beta b_s}{\xi}$$
$$\xi = d(1-(1-d)\beta)b_{aa} + (1-d)^2 \beta b_{ss} - (1-d)(1-(1-2d)\beta)b_{as}$$

Equation (17) shows that long-run equilibrium consumption a^* is not a function of the conformity parameter b_g. It is clear that the effect on long-run consumption of a

[12] See Appendix A for details.

change in price, $\partial a^*/\partial p$, is also not a function of b_g.[13]

Although conformity does not affect the average level of consumption, it does affect the variance. Recall that all the parameters in equation (12) are functions of b_g. For example, an increase in b_g reduces the magnitude of π_1, the direct effect of price on consumption. In fact, π_1 approaches 0 as b_g approaches infinity. Figure 1 illustrates the effect of social interactions on the distribution of consumption using simulations described in the appendix. Comparing Figure 1a to Figure 1b shows that conformity compresses the distribution of consumption relative to a setting with no conformity.

Define the long-run effect of a change in price to be the change in quantity demanded in response to a permanent change in price in all periods. Suppose an individual reaches an expected steady-state level of consumption $a^* = S^*d/(1-d)$. Then this effect is equal to

$$\frac{\partial a^*}{\partial p} = \frac{\pi_1 + \pi_2}{1 - \alpha_1 - \alpha_2 (1-d)/d - \gamma_1 - \gamma_2} \tag{18}$$

Stable dynamics requires that the combined strength of addiction and social interactions not be "too strong". In particular, steady-state stability requires that

$$\alpha_1 + \alpha_2 (1-d)/d + \gamma_1 + \gamma_2 < 1$$

2.2 Proportional spillovers

This subsection considers the case where social interactions take the form of proportional spillovers:

$$G(a_t, E_t[\overline{a}_t]) = b_g a_t E_t[\overline{a}_t]$$

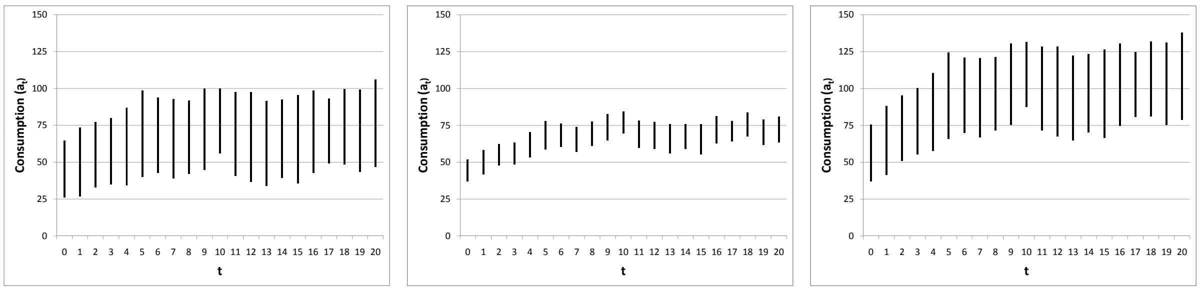
As in Section 2.1, the myopic case is solved first to serve as a baseline:

$$a_t = \alpha'_m S_t + \gamma'_m \overline{a}_t + \pi'_m p_t + \delta'_m x_t + k'_m \tag{19}$$

[13] A more general setting allows p_t and x_t to vary over individuals and time (e.g., Figure 1). In this case a stationary equilibrium results where $E[a_t^*] = E[\overline{a}_t]$. Conformity still has no effect on the average consumption level in this setting, even if the distributions of x_t and p_t are asymmetric.

Figure 1: The effect of social interactions on consumption

(a) No social interactions **(b)** Conformity **(c)** Spillovers



These graphs display the results of three simulations. The vertical lines display the range (minimum to maximum) of consumption for a group of consumers for each period t. Panel (a) displays consumption for an addictive good that exhibits no social interactions. Panel (b) adds conformity to the simulation. Panel (c) adds spillovers instead of conformity to the simulation. See Appendix B for more details.

where

$$\alpha'_m = \frac{b_{as}}{b_{aa}} > 0$$
$$\gamma'_m = \frac{b_g}{b_{aa}} > 0$$
$$\pi'_m = \frac{-\lambda}{b_{aa}} < 0$$
$$\delta'_m = \frac{b_{ax}}{b_{aa}} \lesseqgtr 0$$
$$k'_m = \frac{b_a}{b_{aa}} > 0$$

As in the case of conformity with myopic consumers, consumption is positively related to the stock of past consumption and to group consumption, negatively related to price, and unaffected by the negative consequences of addiction (as captured by b_s and b_{ss}) or by any future determinants of demand. Unlike in the case of conformity, only the coefficient on mean consumption (γ'_m) is a function of the strength of social interactions (b_g).

Relaxing the assumption that consumers are myopic yields the forward-looking demand equation:

$$a_t = \alpha'_1 S_t + \alpha'_2 a_{t+1} + \gamma'_1 \overline{a}_t + \gamma'_2 \overline{a}_{t+1} + \pi'_1 p_t + \pi'_2 p_{t+1} + \delta'_1 x_t + \delta'_2 x_{t+1} + k' \tag{20}$$

where

$$\begin{aligned}
\alpha'_1 &= \frac{b_{as} - (1-d)^2 \beta (b_{as} + b_{ss})}{\Delta'} \\
\alpha'_2 &= (1-d) \beta \frac{b_{aa} + b_{as}}{\Delta'} > 0 \\
\gamma'_1 &= \frac{b_g}{\Delta'} > 0 \\
\gamma'_2 &= -(1-d) \beta \frac{b_g}{\Delta'} < 0 \\
\pi'_1 &= -\frac{\lambda}{\Delta'} < 0 \\
\pi'_2 &= (1-d) \beta \frac{\lambda}{\Delta'} > 0 \\
\delta'_1 &= \frac{b_{ax}}{\Delta'} \lesseqgtr 0 \\
\delta'_2 &= (1-d) \beta \frac{b_{sx} - b_{ax}}{\Delta'} \lesseqgtr 0 \\
k' &= \frac{b_a - (1-d) \beta (b_a - b_s)}{\Delta'} \lesseqgtr 0 \\
\Delta' &= b_{aa} + (1-d)^2 \beta (b_{as} + b_{ss}) > 0
\end{aligned}$$

The signs of the coefficients in (20) are the same as the ones for the conformity demand equation (12). Under proportional spillovers, however, only the coefficients on mean consumption (γ'_1 and γ'_2) are affected by the social interactions parameter b_g. Direct examination of equation (20) reveals that $\partial a_t / \partial b_g > 0$, i.e., an increase in the strength of proportional spillovers increases consumption. This is illustrated in Figure 1c.

The formula for calculating the long-run effect of a change in price is identical to the conformity case and given by (18). The solution to the linear difference equation (20) has the same form as (16). The two solutions differ only in the magnitudes of their coefficients. See Appendix A for details.

3 Identification

There are two main identification concerns associated with estimating the demand equations derived in Section 2. First, the stock of past consumption is necessarily correlated with the error term in the presence of serial correlation. Second, mean group consumption is also correlated with the error term if unobserved shocks to demand are correlated across individuals within the same group. The traditional solution to the first problem is to instrument for stock with lags and, if consumers are forward looking, leads of exogenous determinants of demand such as price (Becker, Grossman, and Murphy 1994; Chaloupka 1991). In this section I show that these instruments can also be used to identify the coefficient on mean group consumption.

Consider a variant of the myopic consumer's demand for an addictive good:

$$a_{igt} = \alpha S_{igt} + \gamma \overline{a}_{gt} + \pi p_{gt} + \delta x_{igt} + \eta_g + e_{igt} \tag{21}$$

where S_{igt} is consumer i's stock of past consumption at time t, $\overline{a}_{gt}$ is the mean consumption for reference group g at time t, and p_{gt} and x_{igt} represent price and individual-level determinants of demand, respectively.[14] The error terms η_g and e_{igt} represent unobserved (to the econometrician) determinants of demand. Equation (21) is a standard "linear-in-means" social interactions model with the addition of the stock term S_{igt} (Manski 1993). There are two central challenges in identifying the endogenous social effect γ when groups are large. First, mean consumption $\overline{a}_{gt}$ is correlated with η_g. Second, estimation is biased if there are group-wide shocks to demand, which causes $Cov\left(\overline{a}_{gt}, e_{igt}\right) \neq 0$.[15]

The dynamics that operate through the stock $S_{igt} = (1-d)\left(S_{igt-1} + a_{igt-1}\right)$ offer a solution to this problem. Addiction implies that individual and group consumption are both related to past determinants of demand.[16] Becker, Grossman, and Murphy (1994) show that this in turn implies that lagged exogenous determinants of demand are valid instruments for S_{igt}. This insight extends naturally to $\overline{a}_{gt}$ in my model.

[14] Equation (21) implicitly assumes that the group-level average of the covariates, $\overline{x}_{gt}$, do not affect demand. This implies, for example, that a consumer's demand may be affected by her education level but her demand is not directly affected by the education level of others in her group. The identification argument presented in this section still holds if one adds "contextual effects" like $\overline{x}_{gt}$ to the demand equation.

[15] Estimation is also biased if individuals choose their reference groups in a way that is correlated with the outcome of interest, which also causes $Cov\left(\overline{a}_{gt}, e_{igt}\right) \neq 0$.

[16] This is confirmed formally by equation (16).

If consumers are forward looking then future values of these variables are also valid instruments.

For example, consider estimating (21) using OLS with fixed effects. Consistency requires $Cov(\overline{a}_{gt}, e_{igt}) = 0$, which rules out the possibility of group-wide shocks to demand. This assumption is unlikely to hold in an empirical analysis, however, because many components of demand are unobserved. If this assumption fails then the econometrician will overestimate the effect of social interactions (γ) because she will incorrectly attribute changes in individual consumption to changes in group consumption when in fact they are due to components in the error term. By contrast, instrumenting for $\overline{a}_{gt}$ with lags of, for example, price identifies γ using only variation in past prices. Validity in this case requires the weaker assumption that these past prices be uncorrelated with the current error term.

Although the strategy outlined above identifies γ (and thus identifies the presence of social interactions), it cannot distinguish between the two forms of social interactions, conformity and proportional spillovers, that were discussed in the previous section. I discuss potential strategies for addressing this in Section 7 but otherwise this topic is beyond the scope of this paper. The next half of this paper applies the identification strategy outlined above to data on smoking consumption.

4 Data

I obtain data on adult smoking behavior from the National Health Interview Survey (NHIS). This cross-sectional, nationally representative survey collects annual information on the health of the civilian noninstitutionalized population of the United States. I analyze thirteen years of the survey between 1970 and 2001.[17] Approximately thirty percent of survey respondents are current smokers. Each survey collected self-reported data on the respondent's current consumption of cigarettes. The 1970, 1978, 1979, and 1980 surveys also collected information on consumption at the time the respondent smoked her greatest average daily quantity, and the 1970 survey additionally collected data on the respondent's consumption in the previous year. The respondent's age, race, gender, marital status, family income, educational attainment, and

[17] 1970, 1978-1980, 1987-1988, 1992, 1995, and 1997-2001. There are other available years with data on smoking behavior but these data lack the respondent's age of smoking initiation and/or the respondent's geographic identifier.

age of smoking initiation are available in each survey. The metropolitan statistical area (MSA) of residence is available for approximately half of all respondents.[18]

One limitation of these data is that, with the exception of the 1970 survey, I do not observe consecutive measures of consumption for individuals. I explain in Section 5 how one can nevertheless use the respondent's reported age of initiation to construct a measure of her stock of past consumption. Unfortunately, I am restricted to estimating a myopic demand equation because I do not observe future consumption.

The factors governing smoking participation have been found to differ from the factors governing consumption conditional on participation so I therefore limit my analysis to current smokers.[19] I define a smoker's reference group to be other smokers who live in her MSA. Some specifications in my analysis consider smaller reference groups by subdividing the MSA along one of the following three demographics: (1) age, classified as forty and under or over forty; (2) gender; and (3) marital status. I limit my analysis to current smokers who live in MSA's with at least 150 other smokers in order to obtain accurate measures of mean group consumption for all possible reference groups. I drop respondents with top-coded ages.[20]

The dependent variable, smoking consumption, is conditional on the respondent being a current smoker. This measure is susceptible to sample selection bias if factors that affect consumption also affect participation. For example, if an increase in the price of cigarettes causes more low intensity smokers to quit than high intensity smokers then the estimated effect of price on consumption will be muted. Fiore et al. (1990) find, however, that daily cigarette consumption does not predict the success of attempted cessation, which suggests that this bias is small.

Table 1 presents summary statistics for the sample. Male smokers consume an average of 17.91 cigarettes per day while female smokers consume 16.00 cigarettes per day.[21] The average male begins smoking around age seventeen and a half, more than one year earlier than the average female. Less than twenty percent of smokers graduated college. Approximately half of the sample is married.

The weighted average price and excise tax per pack of cigarettes are obtained from Orzechowski and Walker (2009) and matched to respondents on the basis of state of

[18]MSA identifiers for individuals from low-population areas are suppressed in the publicly available data in order to ensure confidentiality.

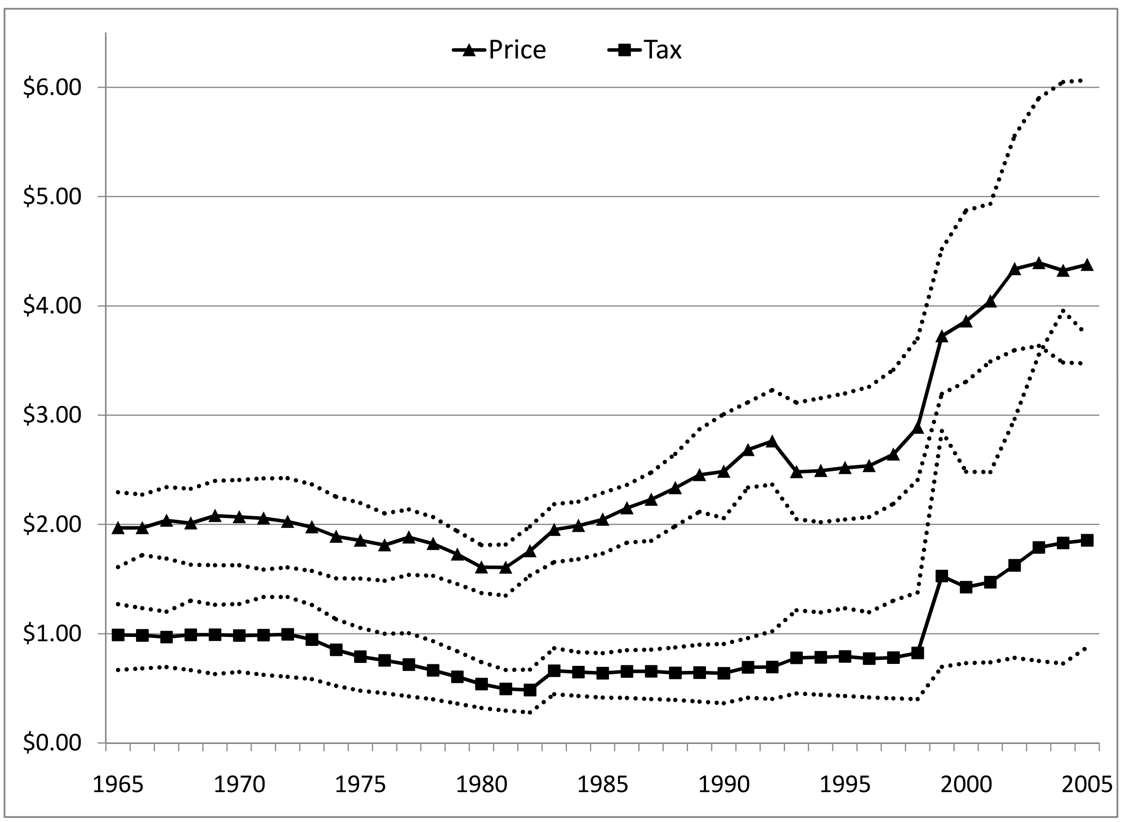
[19]See Jones (1989) and Labeaga (1999).

[20]This only affects respondents who are over the age of 84.

[21]One pack of cigarettes typically includes twenty cigarettes.

residence.[22] I average these data across bordering states for MSA's located across or on state borders. For example, respondents living in Portland, Oregon are matched to the average price for Oregon and Washington. Figure 2 displays prices and taxes for the period 1965-2005 for all fifty states. Prices fell in real terms between 1965 and 1980. They have steadily increased since then with the exception of a price cut in the early 1990's. The sharp rise in 1999 is due to the Master Settlement Agreement between the major cigarette firms and the US state attorney generals. This agreement stipulated that the four largest US cigarette firms pay a minimum of $206 billion to the states over a period of twenty-five years in exchange for a cap on litigation costs.[23]

Figure 2: Real (2009 dollars) price and tax per pack of cigarettes

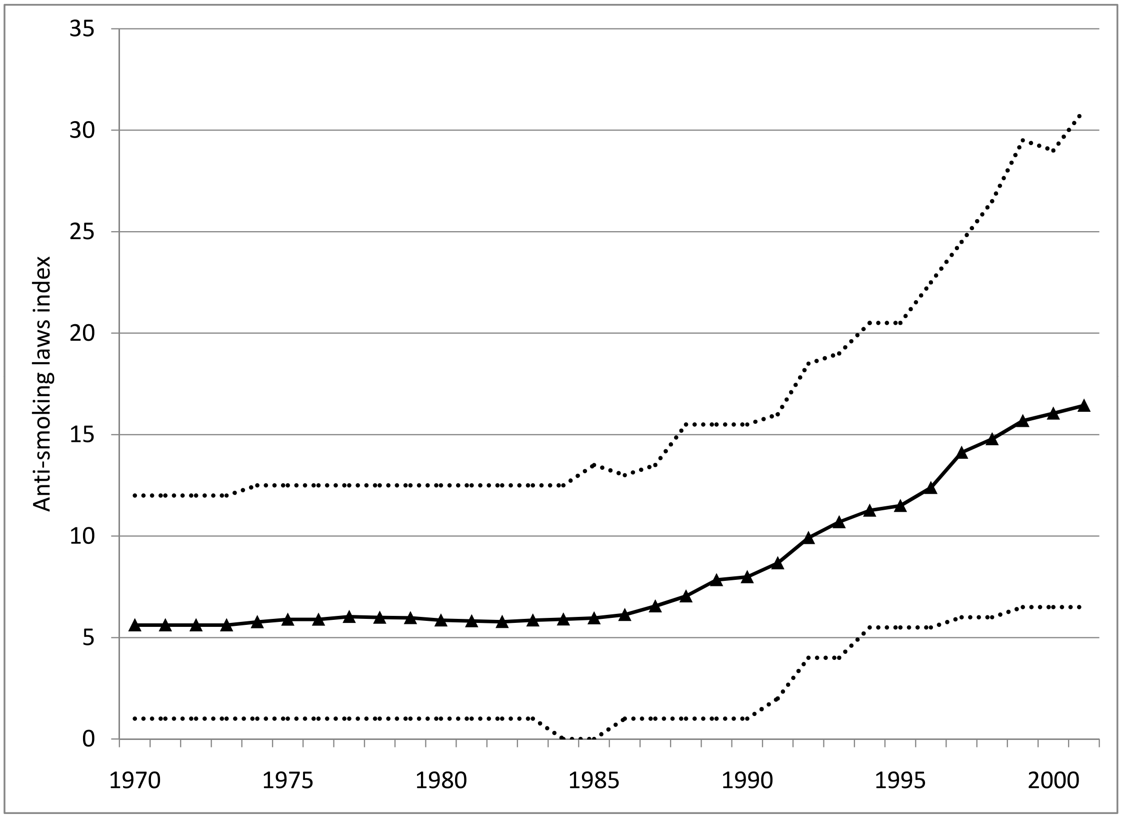


State price and tax data are obtained from Orzechowski and Walker (2009). Prices include taxes. The data are adjusted for inflation using the Urban Consumer Price Index. The solid lines display the national means and the dotted lines display the fifth and ninety-fifth percentiles.

[22] All prices used in the analysis exclude sales tax but include all tobacco-related excise taxes. See Appendix C for details.

[23] See Viscusi and Hersch (2010) for a discussion of the effects of the Master Settlement Agreement on cigarette prices.

Figure 3: Index of state anti-smoking laws



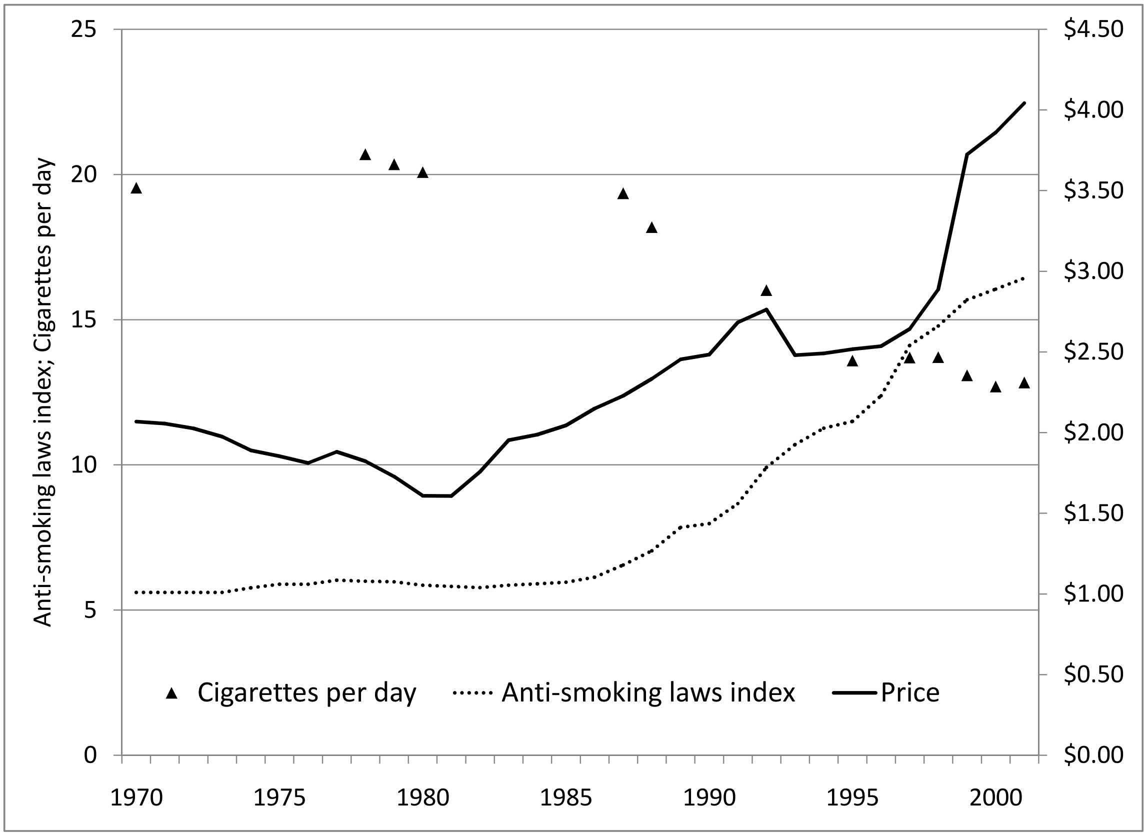
Source: Gruber and Zinman (2001) and ImpacTeen. The index measures the strength of a state's anti-smoking laws. Higher index values indicate states with more restrictive laws. The index ranges from 0 to a possible maximum of 51. The solid line displays the national mean and the dotted lines display the fifth and ninety-fifth percentiles.

Many states enacted various anti-smoking laws during the time period of this analysis. These laws, among other things, ban or restrict the location of cigarette vending machines, outlaw the distribution of free cigarette samples, and regulate cigarette advertising. The National Cancer Institute has developed a set of criteria that measure the strength of these laws (Alciati et al. 1998). These criteria provide the basis for an anti-smoking laws index developed in Gruber and Zinman (2001).[24] The index has a possible range of 0-51 and covers the time period 1973-1998. I extend it to 2001 by incorporating publicly available data from ImpacTeen. See Appendix C for details.

Figure 3 displays the anti-smoking laws index I use in my analysis. I match these

[24]I am grateful to Jonathan Gruber for sharing these data with me.

Figure 4: Number of cigarettes smoked per day, anti-smoking laws, and cigarette prices



This graph plots the average reported number of cigarettes smoked per day in the NHIS sample. It also displays the average cigarette price and the average index of anti-smoking laws. Cigarette prices are in 2009 dollars and include taxes. The anti-smoking laws index ranges in value from 0 to a possible maximum of 51.

data to survey respondents on the basis of their state of residence and year surveyed. Index data are not available for respondents surveyed in 1970. Because there is little to no change in anti-smoking laws in the early 1970's, I match respondents from the 1970 survey to the 1973 index. I average the index data across border states for MSA's located across or on state borders in the same way I did with prices and taxes. The average index for smokers in the sample is 9.59 with a standard deviation of 7.15.

Figure 4 shows that the average number of cigarettes smoked per day began declining in 1980 and that this decline is associated with a corresponding increase in the real price of cigarettes and in the strength of anti-smoking laws.

5 Empirical strategy

Estimating demand requires constructing a measure of a smoker's stock of past consumption, defined as

$$S_t = \sum_{\tau=0}^{t-1} (1-d)^{t-\tau} a_\tau$$

where d is the depreciation rate of the stock, a_τ is the smoker's consumption in period τ, and t is the number of years she smoked. I assume $S_0 = 0$. I do not observe the depreciation rate but medical studies indicate that the effects of addiction typically do not last more than a few years after cessation, which suggests the rate is not low (Fiore 2008, p. 60). I follow Chaloupka (1991) and set the depreciation rate d equal to 0.8 in my analysis.[25] The number of years a respondent has smoked can be calculated by subtracting her age of initiation from her age.

I do not observe a respondent's entire smoking history in any of the NHIS surveys. Instead, I estimate her stock as

$$\widehat{S}_t = \sum_{\tau=0}^{t-1} (1-d)^{t-\tau} \widehat{a}_\tau \tag{22}$$

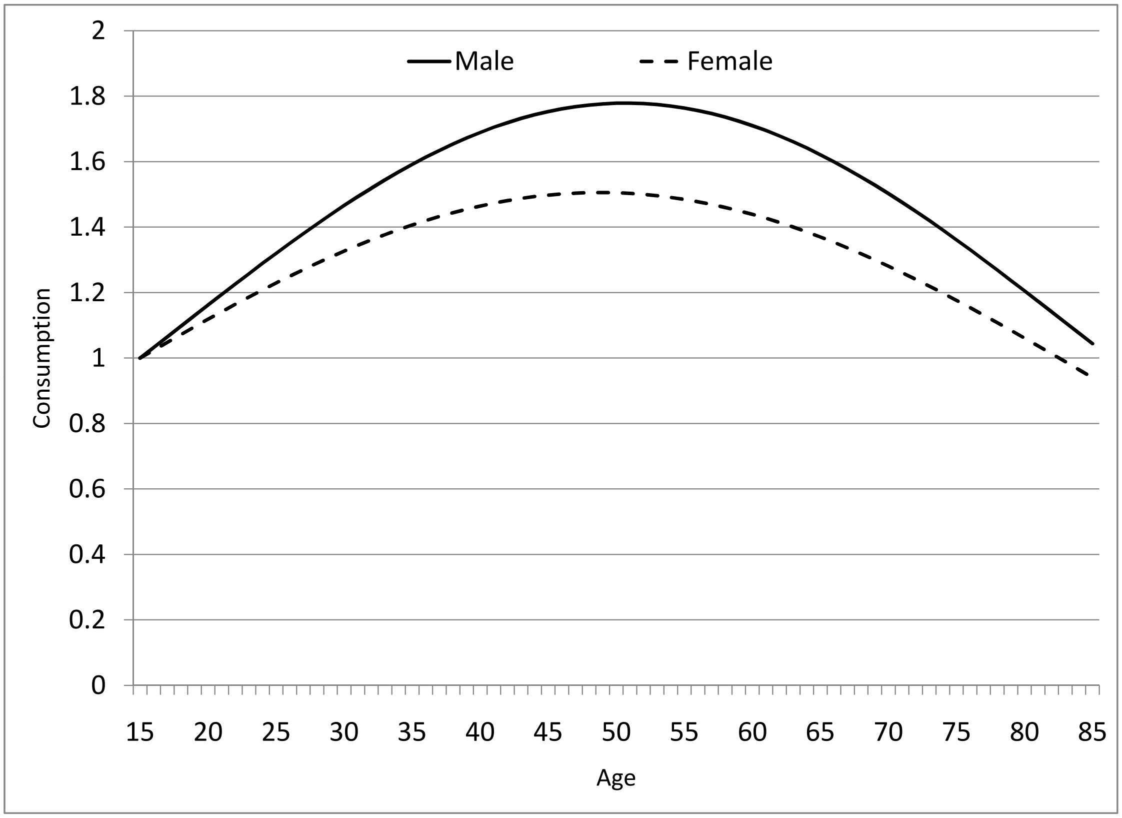
where $\widehat{a}_\tau$ is an estimate of the individual's past smoking consumption. I estimate $\widehat{a}_\tau$ by modeling the age profile of consumption as follows. First, I use the NHIS data to estimate smoking consumption as a function of gender, age, and age-squared. The results of this estimation procedure are displayed in Figure 5 for an 85-year-old smoker who began smoking at age 15. The model captures two key features of smoking behavior: (1) consumption follows an inverted U-shaped curve over an individual's lifetime; and (2) the inverted U-shaped curve is more pronounced for males than females. This curve is then individually scaled for each survey respondent by incorporating data on her smoking history. Data on a_t, a_{t-1}, and $a_{\max} \equiv \max\{a_0, a_1, ..., a_t\}$ are available depending on the year of the survey. If a_{t-1} is available, the curve is scaled so that $\widehat{a}_{t-1} = a_{t-1}$. Otherwise, the curve is scaled so that $\widehat{a}_{\max} = a_{\max}$. If both a_{t-1} and $a_{\max}$ are missing, the curve is scaled so that $\widehat{a}_t = a_t$.

Chaloupka (1991) performs a similar procedure to estimate the stock of past consumption. The main difference between our methods is that his method assumes

[25] Chaloupka (1991) considers a range of depreciation rates centered around 0.8; my results are robust to alternative rates.

$Cov[(1-d)^{t-\tau}, a_\tau] = 0$, which is equivalent to assuming that the curves in Figure 5 are flat. This allows him to set a_τ equal to an individual's mean consumption. My method improves upon his procedure by additionally modeling the age profile of consumption.[26]

Figure 5: Male and female smoking profiles



This graph displays estimated cigarette consumption for an 85-year-old smoker who began smoking at age 15. Consumption at age 15 is normalized to 1. The profile curves are estimated using data on current smokers from the 1970, 1978-1980, 1987-1988, 1992, 1995, and 1997-2001 National Health Interview Surveys.

I estimate a model of demand where smokers are not forward looking:

$$a_{igt} = \alpha\widehat{S}_{igt} + \gamma\overline{a}_{gt} + \pi p_{mt} + \theta\text{Log}[INDEX_{mt}] + x'_{igt}\delta + \eta_m + \eta_t + e_{igt} \qquad (23)$$

$\widehat{S}_{igt}$ is smoker i's estimated stock of past consumption at time t; $\overline{a}_{gt}$ is the mean consumption in her reference group g; p_{mt} is the average price of cigarettes in her

[26]Nevertheless, both methods produce similar results. This is because most of the variation in an individual's stock of past consumption is driven by t, the (known) number of years she smoked.

MSA; $INDEX_{mt}$ is an index of anti-smoking laws in her MSA; and x_{igt} controls for gender, race, age and its square, marital status, log of income, and whether or not the individual is a college graduate. All estimates include MSA and year fixed effects (η_m and η_t).

OLS estimates of (23) are inconsistent because the stock $\widehat{S}_{igt}$ is a function of past consumption.[27] In addition, group consumption $\overline{a}_{gt}$ is correlated with the error term in the presence of unobserved group-wide shocks. Following the identification strategy outlined in Section 3, I account for this by instrumenting for both stock and group consumption with four lags of price and four lags of the log of the anti-smoking laws index.

The long-run price elasticity of demand gives the percentage change in quantity demanded in response to a permanent change in price in all periods. Suppose an individual reaches an expected steady-state level of consumption $a^* = S^*d/(1-d)$. Then the long-run price elasticity is equal to

$$\frac{\partial a^*}{\partial p}\frac{p}{a^*} = \frac{\pi}{1-\alpha\left(1-d\right)/d-\gamma}\frac{p}{a^*}$$

where p and a^* are evaluated at the mean price and consumption levels for the sample. The long-run elasticity with respect to the anti-smoking laws index is calculated similarly. Steady-state stability requires that $\alpha\left(1-d\right)/d+\gamma<1$.

6 Results and discussion

Table 2 reports estimation results for demand equation (23). Column 1 reports a baseline OLS estimate that accounts for neither social interactions nor addiction. Price and the index of anti-smoking laws are negatively related to cigarette consumption but not statistically significant. Females, married individuals, and college graduates are all associated with lower cigarette consumption. Whites consume more cigarettes than non-whites, which is consistent with the literature.[28]

Column 2 adds group consumption as a regressor and thus estimates a static linear-in-means model. These estimates are biased because this regression does not account for addiction or for the possibility that unobserved shocks could be correlated

[27] Serial correlation in the error term causes $\widehat{S}_{igt}$ to be correlated with e_{igt}.
[28] See Gilleskie and Strumpf (2005) and Gruber and Zinman (2001).

across smokers living in the same MSA. Column 3 controls for an individual's stock of past consumption but fails to account for the identification issues discussed in Section 3.

Column 4 of Table 2 addresses the endogeneity of stock by instrumenting for it with four lags of price and four lags of the anti-smoking laws index. The coefficient on stock decreases but remains significant while the coefficient on group consumption increases. As discussed in Section 3, the coefficient on group consumption is biased upwards in the presence of unobserved group-wide shocks. Column 5 addresses this concern by instrumenting for group consumption in addition to stock and thus presents the first consistent estimate of demand. The coefficient on group consumption is 0.246, less than one-half the magnitude of the estimate in Column 4.

The coefficients on price are small and generally insignificant. For example, Column 1 of Table 2 estimates a price elasticity of -0.145 (not shown), which means a 10-percent increase in the price is predicted to reduce consumption by 1.45 percent. The long-run price elasticity estimated by Column 5 is -0.126 (not shown). These estimates are consistent with prior studies of cigarette smoking that employ individual-level data and account for anti-smoking laws.[29] It is also consistent with Chaloupka (1991), who estimates a rational addiction model and obtains insignificant price elasticities for current smokers. My estimates are significantly smaller, however, than the estimates from Becker, Grossman, and Murphy (1994), who apply a rational addiction model to aggregate sales data and estimate a long-run price elasticity of -0.7.

Column 1 estimates an anti-smoking laws index elasticity of -0.046 (not shown) while the correctly specified model of demand presented in Column 5 estimates an elasticity of -0.138 (not shown), which is three times larger. This suggests that anti-smoking laws have a significant impact on smoking intensity, although these results must be interpreted cautiously because anti-smoking laws may be correlated with underlying state-specific trends in smoking behavior. For example, if states with large populations of people who harbor anti-smoking sentiments are more likely to pass anti-smoking laws than other states, then the causal effect of anti-smoking laws

[29] Gruber and Zinman (2001) analyze Monitoring the Future data, which covers the period 1991-1997, and estimate a price elasticity of -0.059 for 12th graders. Wasserman, Manning, Newhouse, and Winkler (1991) analyze data from the National Health Interview Survey and estimate price elasticities that range from 0.06 to -0.283 for the period 1970-1988.

on consumption would be smaller than what these regressions estimate.[30]

The stock of past consumption has a positive and statistically significant effect on consumption, which is consistent with the hypothesis that cigarette smoking is addictive. The coefficient on group consumption is positive and significant, indicating that social interactions have a strong effect on individual consumption. The interpretation of this coefficient depends greatly on the assumed form of social interactions. Under conformity, these results are consistent with a reduction of approximately forty percent in the within-group standard deviation of smoking consumption as compared to a group with no conformity.[31] Under proportional spillovers, these results imply that an increase of one cigarette per day in a group's consumption of cigarettes causes an increase in consumption of 0.246 cigarettes per day for a current smoker.

6.1 Robustness checks

Table 3 reports results when the reference group differs by demography as well as MSA. Column 1 replicates the results from Column 5 of Table 2. Columns 2-4 report results when the reference group differs by age (forty and under or over forty), marital status, and gender, respectively. Changing the reference group does not affect estimates much.

If current cigarette prices are correlated with the error term then the estimates presented so far are inconsistent. The inclusion of MSA and time fixed effects along with the anti-smoking laws index increases the likelihood of price exogeneity. Gruber and Koszegi (2001) and Gruber and Frakes (2006) worry, however, that cigarette firms may set state-specific prices in response to unobserved, time-varying determinants of demand. They argue that cigarette taxes are a better source of exogenous variation than cigarette prices.[32] Table 4 accounts for this possibility by additionally instrumenting for the current price with the current tax. Results are not significantly

[30]DeCicca et al. (2008) provide evidence that accounting for anti-smoking sentiment reduces the effect of price, but not anti-smoking laws, on consumption.

[31]The estimate of forty percent is obtained from numerical simulation and requires making assumptions regarding the distribution of individual determinants of demand. See Appendix B for details.

[32]Chou, Grossman, and Saffer (2006) counter this argument by deriving a model of competition that is consistent with features of the cigarette industry and in which pricing does not depend on the disturbance terms in cigarette demand; moreover, they note that using taxes instead of prices discards useful variation due to state differences in transportation costs, retailing costs, and local competition.

affected.

Finally, Table 5 presents estimates for a model that assumes the depreciation rate of the stock is equal to 0.6 rather than 0.8. The coefficient on stock decreases, as expected, and the coefficient on group consumption becomes insignificant when the reference group corresponds to married individuals. Otherwise, results are not significantly affected.

Examining Tables 3-5 reveals that the largest and most significant estimates for the effect of group consumption on individual consumption correspond to reference groups defined by age. This suggests that social interactions matter the most for individuals within the same age group.

7 Conclusion

This paper develops a general model of demand for a good that exhibits both addiction and social interactions. It proposes a new method for identifying social interactions using the dynamics that arise from addiction. This method provides consistent estimates even in the presence of unobserved group-wide shocks to demand, so long as those shocks are not correlated with past determinants of demand. Modeling the individual decision problem reveals, however, that identifying the presence of social interactions is not sufficient to infer the nature of the externalities generated by social interactions. For example, conformity does not affect the average level of consumption while proportional spillovers increases that level.

I apply my proposed identification methodology to data on cigarette consumption and find strong evidence that group consumption affects individual consumption. Although previous studies have examined the effect of group smoking prevalence on individual smoking participation, this paper is the first to examine how group smoking intensity affects individual smoking intensity. My analysis also suggests that smoking consumption may be strongly affected by anti-smoking laws. This estimated effect increases substantially when I account for both addiction and social interactions.

While I have identified the presence of social interactions, like most studies on social interactions I cannot identify their exact form. One possibility is to pursue a complementary analysis that examines the variance of mean consumption across groups.[33] This method requires making assumptions about the variance-covariance

[33] See Glaeser, Sacerdote, and Scheinkman (1996) and Graham (2008) for details.

structure of the unobservables but in return may allow the researcher to infer their form.

8 Tables

Table 1: NHIS summary statistics

	Male smokers	Female smokers
Cigarettes per day	17.91	16.00
	(12.59)	(11.34)
Age of initiation	17.58	18.80
	(4.72)	(5.71)
Age	40.38	40.88
	(14.32)	(14.86)
White	0.78	0.79
	(0.41)	(0.40)
College graduate	0.16	0.13
	(0.37)	(0.33)
Married	0.56	0.50
	(0.50)	(0.50)
Sample size	8,920	11,051

Standard deviations are given in parentheses. Summary statistics are for the thirteen NHIS surveys from 1970-2001 that contain data on smoking behavior. The sample consists of respondents aged 18 and older who identify themselves as current smokers and have an identifiable metropolitan statistical area of residence. Respondents with top-coded ages or who belong to an MSA with less than 150 surveyed smokers are omitted.

Table 2: OLS and IV estimates of demand

	(1)	(2)	(3)	(4)	(5)
Stock (S_t)			3.370***	2.080***	2.706***
			(0.068)	(0.391)	(0.539)
Group consumption ($\overline{a}_t$)		1.034***	0.244***	0.547***	0.246**
		(0.053)	(0.043)	(0.090)	(0.114)
Price	-0.902	-0.676**	0.133*	-0.177	-0.060
	(0.913)	(0.297)	(0.077)	(0.152)	(0.098)
Log index	-0.777	0.090	-0.087	-0.019	-0.181***
	(0.651)	(0.060)	(0.058)	(0.043)	(0.060)
Female	-2.510***	-2.511***	-0.234**	-1.106***	-0.683
	(0.392)	(0.391)	(0.099)	(0.310)	(0.441)
White	5.709***	5.620***	0.580***	2.510***	1.587*
	(0.474)	(0.479)	(0.135)	(0.599)	(0.873)
Married	-1.118***	-1.123***	-0.344***	-0.642***	-0.497***
	(0.264)	(0.260)	(0.106)	(0.203)	(0.160)
College graduate	-2.329***	-2.331***	-0.591***	-1.257***	-0.934**
	(0.501)	(0.501)	(0.139)	(0.306)	(0.383)
Log income	0.225	0.224	0.113**	0.156	0.135**
	(0.188)	(0.184)	(0.044)	(0.096)	(0.064)
Age	0.732***	0.733***	0.018	0.292***	0.159
	(0.041)	(0.040)	(0.015)	(0.082)	(0.121)
Age-squared	-0.007***	-0.007***	-0.000	-0.003***	-0.002
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
IV for Stock	No	No	No	Yes	Yes
IV for Group consumption	No	No	No	No	Yes
Observations	19,971	19,971	19,971	19,971	19,971
Hansen test (p-value)	N/A	N/A	N/A	0.62	0.62
First-stage F	N/A	N/A	N/A	5.29	27.99

Dependent variable is reported number of cigarettes smoked per day. Standard errors are clustered by MSA and given in parentheses. A */**/*** next to the coefficient indicates significance at the 10/5/1% level. Null hypothesis of Hansen test is that the excluded instruments are uncorrelated with the error term. First-stage F statistic is robust to arbitrary heteroskedasticity and autocorrelation. All regressions include MSA and year fixed effects. The reference group is other smokers in the same MSA.

Table 3: Demand estimates when assuming different reference groups

	(1)	(2)	(3)	(4)
Stock (S_t)	2.706***	2.577***	2.648***	2.822***
	(0.539)	(0.555)	(0.616)	(0.537)
Group consumption ($\overline{a}_t$)	0.246**	0.285**	0.263*	0.204*
	(0.114)	(0.113)	(0.136)	(0.112)
Price	-0.060	-0.111	-0.063	0.006
	(0.098)	(0.109)	(0.104)	(0.102)
Log index	-0.181***	-0.176***	-0.176***	-0.198***
	(0.060)	(0.048)	(0.061)	(0.069)
	[-0.138*]	[-0.146*]	[-0.138*]	[-0.129**]
Female	-0.683	-0.772*	-0.723	-0.136
	(0.441)	(0.455)	(0.499)	(0.172)
White	1.587*	1.774*	1.658*	1.405*
	(0.873)	(0.908)	(0.991)	(0.851)
Married	-0.497***	-0.500***	-0.865***	-0.457***
	(0.160)	(0.172)	(0.310)	(0.158)
College graduate	-0.934**	-0.986***	-0.961**	-0.876**
	(0.383)	(0.375)	(0.429)	(0.357)
Log income	0.135**	0.129*	0.135**	0.129**
	(0.064)	(0.072)	(0.063)	(0.061)
Age	0.159	0.136	0.169	0.134
	(0.121)	(0.105)	(0.138)	(0.119)
Age-squared	-0.002	-0.001	-0.002	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Group	MSA	MSA/age	MSA/married	MSA/gender
IV for Stock	Yes	Yes	Yes	Yes
IV for Group consumption	Yes	Yes	Yes	Yes
Observations	19,971	19,971	19,971	19,971
Hansen test (p-value)	0.62	0.72	0.60	0.60
First-stage F	27.99	11.48	36.29	5.53

Dependent variable is reported number of cigarettes smoked per day. Standard errors are clustered by MSA and given in parentheses. A */**/*** next to the coefficient indicates significance at the 10/5/1% level. Null hypothesis of Hansen test is that the excluded instruments are uncorrelated with the error term. First-stage F statistic is robust to arbitrary heteroskedasticity and autocorrelation. All regressions include MSA and year fixed effects. Long-run elasticities given in brackets. Standard errors for elasticities calculated using delta method.

Table 4: Demand estimates when instrumenting for current price with current tax

	(1)	(2)	(3)	(4)
Stock (S_t)	2.697***	2.590***	2.646***	2.834***
	(0.544)	(0.581)	(0.619)	(0.556)
Group consumption ($\overline{a}_t$)	0.249**	0.281**	0.264*	0.200*
	(0.114)	(0.121)	(0.136)	(0.116)
Price	-0.062	-0.095	-0.058	0.026
	(0.140)	(0.164)	(0.146)	(0.140)
Log index	-0.180***	-0.177***	-0.176***	-0.200***
	(0.060)	(0.049)	(0.061)	(0.069)
	[-0.138*]	[-0.146*]	[-0.138*]	[-0.128**]
Female	-0.689	-0.763	-0.725	-0.138
	(0.442)	(0.467)	(0.500)	(0.171)
White	1.600*	1.755*	1.662*	1.389
	(0.877)	(0.938)	(0.994)	(0.876)
Married	-0.499***	-0.497***	-0.866***	-0.454***
	(0.165)	(0.181)	(0.313)	(0.166)
College graduate	-0.938**	-0.979***	-0.962**	-0.870**
	(0.380)	(0.373)	(0.426)	(0.355)
Log income	0.135**	0.129*	0.135**	0.129**
	(0.065)	(0.073)	(0.063)	(0.062)
Age	0.161	0.134	0.170	0.132
	(0.122)	(0.108)	(0.138)	(0.123)
Age-squared	-0.002	-0.001	-0.002	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Group	MSA	MSA/age	MSA/married	MSA/gender
IV for Stock	Yes	Yes	Yes	Yes
IV for Group consumption	Yes	Yes	Yes	Yes
IV for Price	Yes	Yes	Yes	Yes
Observations	19,971	19,971	19,971	19,971
Hansen test (p-value)	0.62	0.72	0.61	0.59
First-stage F	23.76	9.04	37.24	4.76

Dependent variable is reported number of cigarettes smoked per day. Standard errors are clustered by MSA and given in parentheses. A */**/*** next to the coefficient indicates significance at the 10/5/1% level. Null hypothesis of Hansen test is that the excluded instruments are uncorrelated with the error term. First-stage F statistic is robust to arbitrary heteroskedasticity and autocorrelation. All regressions include MSA and year fixed effects. Long-run elasticities given in brackets. Standard errors for elasticities calculated using delta method.

Table 5: Demand estimates when assuming 60% depreciation instead of 80%

	(1)	(2)	(3)	(4)
Stock (S_t)	1.040***	0.992***	1.022***	1.086***
	(0.209)	(0.214)	(0.241)	(0.207)
Group consumption ($\overline{a}_t$)	0.236*	0.273**	0.249*	0.192
	(0.122)	(0.119)	(0.145)	(0.118)
Price	-0.059	-0.108	-0.059	0.006
	(0.095)	(0.111)	(0.102)	(0.098)
Log index	-0.185***	-0.180***	-0.181***	-0.202***
	(0.058)	(0.047)	(0.059)	(0.068)
	[-0.153*]	[-0.162*]	[-0.153*]	[-0.142**]
Female	-0.665	-0.751	-0.698	-0.142
	(0.451)	(0.464)	(0.512)	(0.167)
White	1.507*	1.690*	1.563	1.317
	(0.897)	(0.929)	(1.023)	(0.870)
Married	-0.513***	-0.515***	-0.859***	-0.474***
	(0.151)	(0.163)	(0.316)	(0.150)
College graduate	-0.923**	-0.974**	-0.945**	-0.863**
	(0.390)	(0.381)	(0.438)	(0.362)
Log income	0.142**	0.136*	0.141**	0.136**
	(0.061)	(0.069)	(0.059)	(0.058)
Age	0.116	0.096	0.124	0.089
	(0.131)	(0.114)	(0.150)	(0.129)
Age-squared	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Group	MSA	MSA/age	MSA/married	MSA/gender
IV for Stock	Yes	Yes	Yes	Yes
IV for Group consumption	Yes	Yes	Yes	Yes
Observations	19,971	19,971	19,971	19,971
Hansen test (p-value)	0.66	0.76	0.63	0.62
First-stage F	23.00	11.05	33.15	5.28

Dependent variable is reported number of cigarettes smoked per day. Standard errors are clustered by MSA and given in parentheses. A */**/*** next to the coefficient indicates significance at the 10/5/1% level. Null hypothesis of Hansen test is that the excluded instruments are uncorrelated with the error term. First-stage F statistic is robust to arbitrary heteroskedasticity and autocorrelation. All regressions include MSA and year fixed effects. Long-run elasticities given in brackets. Standard errors for elasticities calculated using delta method.

A Mathematical appendix

Let private utility take the quadratic form (5). Then maximizing (2) with respect to c_t and subject to the budget constraint (3) yields

$$c_t = \frac{-\lambda + u_c + u_{ac}a_t + u_{sc}S_t + u_{xc}x_t}{u_{cc}} \tag{24}$$

where λ is the marginal utility of wealth. Plugging this result into (5) allows individual utility (4) to be expressed as

$$\begin{aligned} V^*(t) = & -\frac{1}{2}\left(b_{aa}a_t^2 + b_{ss}S_t^2 + b_{xx}x_t^2\right) + b_{as}a_tS_t + b_{ax}a_tx_t + b_{sx}S_tx_t \\ & + b_a a_t + b_s S_t + b_x x_t + G\left(a_t, E_t\left[\overline{a}_t\right]\right) \end{aligned} \tag{25}$$

where

$$\begin{aligned} b_{aa} &= \frac{u_{aa}u_{cc} - u_{ac}^2}{u_{cc}} > 0 \\ b_{ss} &= \frac{u_{ss}u_{cc} - u_{sc}^2}{u_{cc}} > 0 \\ b_{xx} &= \frac{u_{xx}u_{cc} - u_{xc}^2}{u_{cc}} > 0 \\ b_{as} &= \frac{u_{ac}u_{sc} + u_{as}u_{cc}}{u_{cc}} > 0 \text{ if } u_{ac}u_{sc} \geq 0 \\ b_{ax} &= \frac{u_{ac}u_{xc} + u_{ax}u_{cc}}{u_{cc}} \lesseqgtr 0 \\ b_{sx} &= \frac{u_{sc}u_{xc} + u_{sx}u_{cc}}{u_{cc}} \lesseqgtr 0 \\ b_a &= \frac{u_c u_{ac} + u_a u_{cc}}{u_{cc}} \lesseqgtr 0 \\ b_s &= \frac{u_c u_{sc} + u_s u_{cc}}{u_{cc}} \lesseqgtr 0 \\ b_x &= \frac{u_{ac}u_{xc} + u_{ax}u_{cc}}{u_{cc}} \lesseqgtr 0 \end{aligned}$$

I suppress the constant term in (25) because it is not needed for maximization.

Maximizing $\sum_{t=1}^{\infty} \beta^{t-1} V^*(t)$ with respect to a_t implies the following first order condition:

$$
\begin{aligned}
0 &= V_a^*(t) - \lambda p_t + \Sigma_{i=1}^{\infty} \beta^i V_s^*(t+i) \frac{\partial S_{t+i}}{\partial a_t} \\
&= V_a^*(t) - \lambda p_t + \beta(1-d) V_s^*(t+1) + \Sigma_{i=2}^{\infty} \beta^i V_s^*(t+i) \frac{\partial S_{t+i}}{\partial a_t}
\end{aligned}
$$

where $V_a^*(t) = \frac{\partial V^*(t)}{\partial a_t}$ and $V_s^*(t) = \frac{\partial V^*(t)}{\partial S_t}$. The corresponding first order condition for a_{t+1} is

$$
\begin{aligned}
0 &= V_a^*(t+1) - \lambda p_{t+1} + \Sigma_{i=1}^{\infty} \beta^i V_s^*(t+1+i) \frac{\partial S_{t+1+i}}{\partial a_{t+1}} \\
&= \beta V_a^*(t+1) - \lambda \beta p_{t+1} + \Sigma_{i=2}^{\infty} \beta^i V_s^*(t+i) \frac{\partial S_{t+i}}{\partial a_{t+1}} \\
&= \beta V_a^*(t+1) - \lambda \beta p_{t+1} + \Sigma_{i=2}^{\infty} \beta^i V_s^*(t+i) \frac{\partial S_{t+i}}{\partial a_t} / (1-d) \\
&= \beta(1-d) V_a^*(t+1) - \lambda \beta (1-d) p_{t+1} + \Sigma_{i=2}^{\infty} \beta^i V_s^*(t+i) \frac{\partial S_{t+i}}{\partial a_t}
\end{aligned}
$$

where I use that $\frac{\partial S_{t+i}}{\partial a_{t+1}} = \frac{\partial S_{t+i}}{\partial a_t} / (1-d)$. Equating the first order conditions for a_t and a_{t+1} yields

$$
V_a^*(t) - \lambda p_t + \beta(1-d) V_s^*(t+1) = \beta(1-d) V_a^*(t+1) - \lambda \beta (1-d) p_{t+1} \quad (26)
$$

Plugging in the quadratic utility form (25) and then the social utility specification (8) or (9) yields equations (12) and (20), respectively.

Demand equation (12) can be equivalently written as

$$
S_{t+2} - \frac{1+\alpha_2(1-d)}{\alpha_2} S_{t+1} + \frac{(1+\alpha_1)(1-d)}{\alpha_2} S_t = -\frac{1-d}{\alpha_2} h_{t+2} \quad (27)
$$

where

$$
h_{t+2} = \gamma_1 \overline{a}_t + \gamma_2 \overline{a}_{t+1} + \pi_1 p_t + \pi_2 p_{t+1} + \delta_1 x_t + \delta_2 x_{t+1} + k
$$

Equation (27) is a second-order linear difference equation that can be solved using standard methods.[34] Rewrite (27) using lag operator notation:

$$
\left(1 - \beta_1 L + \beta_2 L^2\right) S_t = -\beta_3 h_t \quad (28)
$$

[34] Sargent (1987) is a good reference.

where

$$\beta_1 = \frac{1 + \alpha_2 (1 - d)}{\alpha_2}$$
$$\beta_2 = \frac{(1 + \alpha_1)(1 - d)}{\alpha_2}$$
$$\beta_3 = \frac{1 - d}{\alpha_2}$$

Factorizing the lag polynomial yields

$$\left(1 - \beta_1 L + \beta_2 L^2\right) = (1 - \phi_1 L)(1 - \phi_2 L)$$

where

$$\phi_1 = \frac{2\beta_2}{\beta_1 - \sqrt{(\beta_1)^2 - 4\beta_2}}$$
$$\phi_2 = \frac{2\beta_2}{\beta_1 + \sqrt{(\beta_1)^2 - 4\beta_2}}$$

The general solution to (28) is

$$S_t = \frac{-\beta_3}{(1 - \phi_1 L)(1 - \phi_2 L)} h_t + c_1 (\phi_1)^t + c_2 (\phi_2)^t \tag{29}$$

where c_1 and c_2 are constants. I set $c_1 = 0$ in order to ensure stability.[35] $\phi_1 \neq \phi_2$ implies the identity

$$\frac{1}{(1 - \phi_1 L)(1 - \phi_2 L)} = \frac{1}{\phi_1 - \phi_2} \left(\frac{\phi_1}{1 - \phi_1 L} - \frac{\phi_2}{1 - \phi_2 L} \right)$$

which is used to rewrite (29) as

$$S_t = \left(\frac{-\beta_3 \phi_1}{(\phi_1 - \phi_2)(1 - \phi_1 L)} + \frac{\beta_3 \phi_2}{(\phi_1 - \phi_2)(1 - \phi_2 L)} \right) h_t + c_2 (\phi_2)^t$$

[35] Stable dynamics also require that $(\beta_1)^2 - 4\beta_2 \geq 0$ and that $\phi_2 < 1$. The former prohibits oscillatory behavior while the latter ensures that consumption does not diverge to infinity. I assume that $\phi_1 > 1$ and $\phi_2 < 1$ for ease of exposition.

Using that $\frac{1}{1-\phi L}x_t = \sum_{j=0}^{\infty}\phi^j x_{t-j}$ (if $\phi < 1$) and $\frac{1}{1-\phi L}x_t = -\sum_{j=1}^{\infty}\phi^{-j} x_{t+j}$ (if $\phi > 1$) allows one to write the solution for S_t as a function of infinite sums:

$$\begin{aligned} S_t &= \frac{\beta_3\phi_1}{\phi_1-\phi_2}\sum_{j=1}^{\infty}(\phi_1)^{-j}h_{t+j} + \frac{\beta_3\phi_2}{\phi_1-\phi_2}\sum_{j=0}^{\infty}(\phi_2)^{j}h_{t-j} + c_2(\phi_2)^t \qquad (30)\\ &= K_1\sum_{j=1}^{\infty}(\phi_1)^{-j}h_{t+j} + K_2\sum_{j=0}^{t-1}(\phi_2)^{j}h_{t-j} + K_2(\phi_2)^t\sum_{j=0}^{\infty}(\phi_2)^{j}h_{-j} + c_2(\phi_2)^t \end{aligned}$$

where

$$\begin{aligned} K_1 &= \beta_3\frac{\phi_1}{\phi_1-\phi_2}\\ K_2 &= \beta_3\frac{\phi_2}{\phi_1-\phi_2} \end{aligned}$$

Solving for the initial condition S_0 yields

$$S_0 = K_1\sum_{j=1}^{\infty}(\phi_1)^{-j}h_j + K_2\sum_{j=0}^{\infty}(\phi_2)^{j}h_{-j} + c_2 \qquad (31)$$

Solving (31) for c_2 and plugging the result into equation (30) yields, after some algebra, the particular solution to (28):

$$S_t = K_1\sum_{j=1}^{\infty}(\phi_1)^{-j}h_{t+j} + K_2\sum_{j=0}^{t-1}(\phi_2)^{j}h_{t-j} + (\phi_2)^t\left(S_0 - K_1\sum_{j=1}^{\infty}(\phi_1)^{-j}h_j\right) \qquad (32)$$

The myopic demand equation (11) is a first-order difference equation. Its solution is

$$S_t = (1-d)\sum_{j=0}^{t-1}(\phi)^{j}h_{t-j} + (\phi)^t S_0 \qquad (33)$$

where

$$\begin{aligned} \phi &= (1-d)(1+\alpha)\\ h_t &= \pi p_{t-1} + \delta x_{t-1} + \gamma\overline{a}_{t-1} \end{aligned}$$

B Simulations appendix

Figure 1 displays the range of consumption for one-hundred individuals from three different simulations. All simulations are generated using equation (32). The parameters used are

$$\begin{aligned}
\beta &= 0.75 \\
d &= 0.5 \\
\lambda &= 1 \\
b_a &= 15 \\
b_s &= 10 \\
b_{as} &= 0.2 \\
b_{aa} &= b_{ss} = b_{xx} = 0.3 \\
b_{ax} &= 2 \\
b_{sx} &= 4
\end{aligned}$$

Prices p_t and covariates x_{it} are drawn i.i.d. $N(10, 1)$ and $N(0, 1)$, respectively. The initial stock of past consumption, S_{i0}, was drawn with a uniform distribution in $[0, 10]$. Data for Figure 1a, which does not include social interactions, are generated directly by equation (32) by setting the parameter $b_g = 0$.[36] Data for Figures 1b and 1c, which include social interactions, require a multi-step procedure:

1. Generate individual consumption data with equation (32) assuming no social interactions (i.e., $b_g = 0$).

2. Calculate mean consumption $\overline{a}'_t$ for each period t.

3. Generate individual consumption data with equation (32) assuming social interactions and using the mean consumption $\overline{a}'_t$ calculated in step 2. I set $b_g = 0.5$ for the conformity specification and $b_g = 0.05$ for the proportional spillovers specification.

4. Calculate mean consumption $\overline{a}''_t$ for each period t using the consumption data generated in step 3. Repeat steps 2-4 until $\left|\overline{a}'_t - \overline{a}''_t\right| < \varepsilon$ where $\varepsilon > 0$ is a

[36] Recall that $h_t = \gamma_1 \overline{a}_t + \gamma_2 \overline{a}_{t+1} + \pi_1 p_t + \pi_2 p_{t+1} + \delta_1 x_t + \delta_2 x_{t+1} + k$ in equation (32). The parameters γ_1 and γ_2, defined in the text of the paper, are equal to 0 when $b_g = 0$.

tolerance parameter set arbitrarily close to 0. I set the tolerance parameter ε to 1×10^{-6} for my simulations.

Equation (33) is used to simulate data to estimate the reduction in variance implied by the estimates reported in Column 5 of Table 2 under the assumption that social interactions take the conformity form (8). The parameters used are

$$\begin{aligned}
\lambda &= 0 \\
d &= 0.8 \\
b_a &= 1 \\
b_{aa} &= 0.3 \\
b_{as} &= 1 \\
b_{ax} &= 0.275 \\
b_g &= 0.09
\end{aligned}$$

The covariates x_{it} are distributed $N(0, 50)$.[37] Data are simulated for fifty time periods. These parameters generate the average consumption level observed in the NHIS surveys and match the regression estimates displayed in Column 5 of Table 2. I account for the effect of anti-smoking laws by including an additional determinant of demand, x_2, whose logarithm is distributed uniformly on $[0, 51]$. I compare the standard deviation of consumption produced by this simulation with the standard deviation of consumption produced by a simulation with no social interactions ($b_g = 0$) using the formula:

$$1 - \frac{Stdev\,[a_{it}]^{b_g=0.25}}{Stdev\,[a_{it}]^{b_g=0}}$$

C Data appendix

Orzechowski and Walker (2009) report annual state-level data on the (tax-inclusive) price of a pack of cigarettes for the period 1954-2009. The reported prices are weighted by sales across all brands, including generics.[38] With the exception of 1958-60 and

[37] The high simulated variance of x_{it} accounts for the large variance in consumption observed in the NHIS surveys. (See Table 1.)

[38] Beginning in 1990, they additionally report the average price excluding generic brands. I do not use that data.

1991, all reported prices are as of November 1. Cigarette taxes are available at the monthly level and include both state and federal cigarette excise taxes. I interpolate the price data and adjust it to account for relevant tax changes to obtain prices as of July 1. For example, the average price for July 1, 1970 is calculated as 4/12 times the reported price (net of tax) on November 1, 1969 plus 8/12 times the reported price (net of tax) on November 1, 1970 plus the reported tax for July 1, 1970.

Neither the prices nor the taxes reported in Orzechowski and Walker (2009) generally include local excise taxes. I account for this by estimating an annual average local excise tax for each state and adding it to the reported prices and taxes. The estimate is achieved by summing the reported county and city tobacco tax revenues for each state and then dividing by state cigarette sales.

The Master Settlement Agreement requires cigarette companies to pay participating states an amount proportional to their sales in those states. Viscusi and Hersch (2010) demonstrate that this payment is an implicit tax. Orzechowski and Walker (2009) report annual settlement payments for each state. This allows me to estimate the effective tax by dividing these payments by the reported sales for each state.

The Anti-Smoking Laws Index is based on the Youth Access Index in Gruber and Zinman (2001). The Youth Access Index covers the years 1973-1998 and is constructed by scoring each state in twelve different categories: minimum age of purchase; packaging; clerk intervention; photo identification; vending-machine availability; free distribution; graduated penalties; random inspections; statewide enforcement; advertising; licensing; and restrictions on minors. Summing across all twelve categories yields the Youth Access Index, which ranges in value from 0 to a possible maximum of 51. See Gruber and Zinman (2001) and Alciati et al. (1998) for additional details.

The Anti-Smoking Laws Index extends the Youth Access Index to include the years 1999-2001 by supplementing it with publicly available data from ImpacTeen.[39] The ImpacTeen data score each state using the same criteria as Alciati et al. (1998). They do not include scores for the advertising, licensing, and restrictions on minors categories. I assume that the scores for these three categories are unchanged since 1998 when constructing the Anti-Smoking Laws Index.

Smoking behavior data are obtained from the 1970, 1978-1980, 1987-1988, 1992,

[39] See http://www.impacteen.org/tobaccodata.htm. These data cover the years 1991-2007. There are occasional discrepancies between the scores reported by ImpacTeen and those reported by Gruber and Zinman (2001). I use the scores from Gruber and Zinman (2001) in those cases.

1995, and 1997-2001 National Health Interview Survey (NHIS). Respondents in surveys prior to 1993 are coded as current smokers if they responded affirmatively to the question “Do you smoke cigarettes now?”. Later respondents are coded as current smokers if they reported smoking either “every day” or “some days”.

The non-response rate is high for the total family income question beginning in 1997. I address this by using the imputed income data files provided by NHIS on their website.

References

Akerlof, G. A. (1991, May). Procrastination and obedience. *American Economic Review 81*(2), 1–19.

Alciati, M., M. Frosh, S. Green, R. Brownson, P. Fisher, R. Hobart, A. Roman, R. Sciandra, and D. Shelton (1998). State laws on youth access to tobacco in the united states: measuring their extensiveness with a new rating system. *Tobacco Control 7*(4), 345–352.

Aristei, D. and L. Pieroni (2009). Addiction, social interactions and gender differences in cigarette consumption. *Empirica 36*(3), 245–272.

Becker, G. S., M. Grossman, and K. M. Murphy (1990, April). An empirical analysis of cigarette addiction. *NBER Working Paper Series* (3322).

Becker, G. S., M. Grossman, and K. M. Murphy (1994, June). An empirical analysis of cigarette addiction. *American Economic Review 84*(3), 396–418.

Becker, G. S. and K. M. Murphy (1988, August). A theory of rational addiction. *Journal of Political Economy 96*(4), 675–700.

Bernheim, B. D. (1994, October). A theory of conformity. *Journal of Political Economy 102*(5), 841–77.

Binder, M. and M. H. Pesaran (2001, January). Life-cycle consumption under social interactions. *Journal of Economic Dynamics and Control 25*(1-2), 35–83.

Bisin, A., U. Horst, and O. Ozgur (2006, March). Rational expectations equilibria of economies with local interactions. *Journal of Economic Theory 127*(1), 74–116.

Blume, L. E., W. A. Brock, S. N. Durlauf, and Y. M. Ioannides (2010, October). Identification of social interactions. Economics Series 260, Institute for Advanced Studies.

Brock, W. A. and S. N. Durlauf (2001, January). Interactions-based models. In J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics*, Volume 5 of *Handbook of Econometrics*, Chapter 54, pp. 3297–3380. Elsevier.

Case, A. C. and L. F. Katz (1991, May). The company you keep: The effects of family and neighborhood on disadvantaged youths. NBER Working Papers 3705, National Bureau of Economic Research, Inc.

Chaloupka, F. (1991, August). Rational addictive behavior and cigarette smoking. *Journal of Political Economy 99*(4), 722–42.

Chou, S., M. Grossman, and H. Saffer (2006). Reply to jonathan gruber and michael frakes. *Journal of Health Economics 25*(2), 389–393.

Christakis, N. and J. Fowler (2007). The spread of obesity in a large social network over 32 years. *New England Journal of Medicine 357*(4), 370–379.

DeCicca, P., D. Kenkel, A. Mathios, Y.-J. Shin, and J.-Y. Lim (2008). Youth smoking, cigarette prices, and anti-smoking sentiment. *Health Economics 17*(6), 733–749.

Durlauf, S. and H. Tanaka (2008). Understanding regression versus variance tests for social interactions. *Economic Inquiry 46*(1), 25–28.

Fiore, M. (2008). *Treating tobacco use and dependence: 2008 update: Clinical practice guideline*. DIANE Publishing.

Fiore, M., T. Novotny, J. Pierce, G. Giovino, E. Hatziandreu, P. Newcomb, T. Surawicz, and R. Davis (1990). Methods used to quit smoking in the united states. *JAMA: The Journal of the American Medical Association 263*(20), 2760.

Gaviria, A. and S. Raphael (2001, May). School-based peer effects and juvenile behavior. *The Review of Economics and Statistics 83*(2), 257–268.

Gilleskie, D. B. and K. S. Strumpf (2005). The behavioral dynamics of youth smoking. *Journal of Human Resources 40*(4), 822–866.

Glaeser, E. L., B. Sacerdote, and J. A. Scheinkman (1996, May). Crime and social interactions. *The Quarterly Journal of Economics 111*(2), 507–48.

Graham, B. (2008). Identifying social interactions through conditional variance restrictions. *Econometrica 76*(3), 643.

Graham, B. and J. Hahn (2005). Identification and estimation of the linear-in-means model of social interactions. *Economics Letters 88*(1), 1–6.

Gruber, J. and M. Frakes (2006). Does falling smoking lead to rising obesity? *Journal of Health Economics 25*(2), 183–197.

Gruber, J. and B. Koszegi (2001). Is Addiction “Rational”? Theory and Evidence. *Quarterly Journal of Economics 116*(4), 1261.

Gruber, J. and J. Zinman (2001). Youth smoking in the united states: Evidence and implications. In *Risky Behavior among Youths: An Economic Analysis*, NBER Chapters, pp. 69–120. National Bureau of Economic Research, Inc.

Jones, A. (1989). A double-hurdle model of cigarette consumption. *Journal of Applied Econometrics 4*(1), 23–39.

Krauth, B. V. (2007, July). Peer and selection effects on youth smoking in california. *Journal of Business & Economic Statistics 25*, 288–298.

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

Levy, M. (2010). An empirical analysis of biases in cigarette addiction.

Manski, C. F. (1993, July). Identification of endogenous social effects: The reflection problem. *Review of Economic Studies 60*(3), 531–42.

Nakajima, R. (2007, 07). Measuring peer effects on youth smoking behaviour. *Review of Economic Studies 74*(3), 897–935.

O'Donoghue, T. and M. Rabin (2002). Addiction and present-biased preferences. *Mimeo*.

Orphanides, A. and D. Zervos (1995, August). Rational addiction with learning and regret. *Journal of Political Economy 103*(4), 739–58.

Orzechowski, W. and R. Walker (2009). *The Tax Burden on Tobacco: Historical Compilation*, Volume 44.

Powell, L. M., J. A. Tauras, and H. Ross (2005, September). The importance of peer effects, cigarette prices and tobacco control policies for youth smoking behavior. *Journal of Health Economics 24*(5), 950–968.

Sacerdote, B. (2001). Peer effects with random assignment: Results for dartmouth roommates*. *Quarterly Journal of Economics 116*(2), 681–704.

Sargent, T. (1987). *Macroeconomic theory*, Volume 2. Academic Press.

The Economist (2011, July 30). Trim staff, fat profits? *The Economist*.

Viscusi, W. K. and J. Hersch (2010, July). Tobacco regulation through litigation: The master settlement agreement. In *Regulation vs. Litigation: Perspectives* *from Economics and Law*, NBER Chapters, pp. 71–101. National Bureau of Economic Research, Inc.

Wasserman, J., W. Manning, J. Newhouse, and J. Winkler (1991). The effects of excise taxes and regulations on cigarette smoking. *Journal of Health Economics 10*(1), 43–64.

Winston, G. C. (1980, December). Addiction and backsliding : A theory of compulsive consumption. *Journal of Economic Behavior & Organization 1*(4), 295–324.