Peers and prices: Explaining the black-white youth smoking gap *

Brian Krauth
Simon Fraser University

August 10, 2001

Abstract

In 1976 black and white teenagers in the United States were about equally likely to be cigarette smokers. By the early 1990's, the smoking rate of black teenagers had dropped to one-third that of white teenagers. This paper analyzes the role of peers, prices, and other factors in explaining this divergence in behavior. I find that the dynamics of youth smoking can best be explained by the combination of rising prices in the 1980's, a higher price elasticity for black teenagers, and the amplifying effects of social interactions (peer effects). In the process, I develop and implement several empirical tools for the analysis of the equilibrium implications of social interactions. In particular, I develop a procedure for determining whether peer influence is strong enough to produce multiple equilibria, and a procedure for estimating the "social multiplier" associated with peer effects. I find that the multiple equilibria explanation is not empirically supported, but that the social multiplier effect is large enough to account for roughly half the difference in smoking rates.

JEL codes: C5, I1
Keywords: Social interactions, peer effects, smoking.

1 Introduction

In 1976 black and white teenagers in the United States were about equally likely to be cigarette smokers. By the early 1990's, the smoking rate of black teens had fallen by almost 70% while the smoking rate of white teens had fallen by only 25%. This dramatic divergence in behavior has attracted significant attention from public health researchers (Faulkner et al. 1996, Nelson et al. 1995, Novotny et al. 1996), but their attempts to find an empirically supported explanation

*I am grateful to Kim-Sau Chung, Steven Durlauf, Judith Giles, Edward Glaeser, and Krishna Pendakur for comments. All errors are mine. Department of Economics, Simon Fraser University, Burnaby BC V5A 1S6. Email:bkrauth@sfu.ca.

have met with limited success. The difficulty encountered by these researchers is that the standard characteristics (parental behavior, income, education) have not exhibited enough variation and more exotic explanations (targeted advertising, shifting social norms) are difficult to assess with the existing data. As a result, the divergence in smoking rates by race remains a puzzle.

This paper evaluates several candidate explanations and finds that only one is supported in the data. I find that the only satisfactory explanation is a combination of three factors - an increase in the real price of cigarettes from 1980 to 1993, a higher price elasticity of participation[1] for black teens, and a multiplier effect caused by the propensity of teens to smoke when their friends smoke.

This conclusion is reached through several steps. To evaluate the explanations most often suggested in the literature, I estimate a baseline logit model with explanatory variables describing family background, disposable income, environmental factors, and prices. Peer behavior is excluded and all coefficients are constrained to be race-invariant. The resulting model is then used to construct predictions of the time series of smoking rates by race, and these predictions are compared to the actual time series depicted in Figure 1. In accordance with the existing literature, the baseline model explains none of the divergence in smoking rates. It predicts almost no difference in the smoking behavior of black and white teenagers, implying that none of the variables in the regression are useful in explaining the observed divergence in smoking rates.

Next, I analyze the impact of social interactions by estimating models in which the smoking rate of the teen's peer group is included as an explanatory variable. A well-known result in the social interactions literature[2] is that when peer behavior has an impact on a person's choices, the equilibrium configuration of choices in a social group may not be unique. For example, if each member of a group is a pure conformist (i.e., will always choose whatever the majority of his friends choose), then smoking rates of zero and 100 percent are both equilibria for the group. In this case rapid changes in group smoking rates can result

[1] The literature distinguishes between participation elasticity (the effect on the probability of smoking) and consumption elasticity (the effect on the amount smoked).

[2] "Social interactions" is simply a general term for situations in which a person's choices vary directly (rather than through the price mechanism) with the choices of some reference group.

from equilibrium-switching or "sunspot" dynamics. I investigate the equilibrium-switching explanation by developing and implementing a test for the consistency of model estimates with multiple equilibria. The results indicate that peer effects are not strong enough to imply multiple equilibria, so the equilibrium-switching explanation can be rejected.

Even with uniqueness of equilibria, peer effects can still play an important role by amplifying the effect of any aggregate shock. For example, suppose that the price of cigarettes increases. This will have both a direct and indirect effect on each individual – it will increase the price faced by that individual, and it will decrease the fraction of friends who smoke. Because of the indirect effect, the price elasticity of a group's smoking rate may be several times the price elasticity of any member of the group. This amplifying effect can be quantified using a "social multiplier". I develop and implement a procedure for estimating the social multiplier, and find that the social multiplier is approximately 2.5. In other words, the elasticity of a group's smoking rate to an aggregate shock is 2.5 times the elasticity of an individual's smoking probability to that same shock. While social interaction effects alone do not explain the divergence in smoking rates, this result indicates they are likely to play an important subsidiary role in the explanation.

Finally, I allow the price response to differ for black and white teenagers. Chaloupka and Pacula (1998) find that black teens have higher price elasticities than white teens, and suggest that this may explain the puzzle. The results presented confirm Chaloupka and Pacula's conjecture. I find that black teens are substantially more responsive to price changes than white teens. Combined with the time series of prices, which rise dramatically in the 1980's and fall slowly in the 1990's, the model's predictions include a divergence in smoking rates which is qualitatively similar to that seen in Figure 1. In the absence of corresponding shifts to peer behavior, the shifts in predicted smoking rates are fairly small. However, once the amplifying effects of peer behavior are included, the predictions match the actual behavior of smoking rates quite well. These results indicate that explaining the large divergence in smoking rates requires the interaction of all three factors - the change in prices, the difference in price response, and the social multiplier.

1.1 Related literature

There is a large economic literature on the determinants of smoking. Becker and Murphy's (1988) "rational addiction" model exemplifies the economic approach to smoking by viewing the decision to smoke as a dynamic optimization problem. There results imply that the price elasticity of young people should be higher than older people (especially if the price shock is permanent), and that smoking should respond to shifts in future income or prices. These results, combined with the greater public support for public intervention in the choices of young people, suggest that policy makers looking to reduce smoking should focus on youth.

Numerous econometric studies have investigated the price elasticity of participation (the decision to smoke or not) in both adults and young people. Both tobacco excise taxes and other tobacco control policies are mostly set at the state level, so cross-state variations in after-tax costs are exploited to estimate elasticities. Although many of these studies face problems controlling for unobserved state characteristics (states with low tobacco taxes and few regulations tend to be tobacco-growing states), most studies have found a price elasticity of participation between -0.5 and -0.7 (Gruber and Zinman 2001). Several studies (Gruber and Zinman 2001, Chaloupka and Pacula 1998) also find substantial variation in price elasticity among teens, especially between white and black teens.

Several studies in the public health literature (Faulkner et al. 1995, van Roosmalen et al. 1989 and 1992, Wang et al. 1995) find that peer smoking is a strong influence on a young person's decision to smoke. In addition, these authors further investigate the mechanisms by which peer behavior influences the smoking decision. They find that few adolescents report that they are pressured to smoke or offered cigarettes by peers, but rather that peer influences affect a teen's estimates of the risks and benefits of smoking. This may reflect either information gained by watching smokers or direct impact on the benefits of smoking provided by being able to smoke in a group.

The econometric literature on social interactions, which includes peer effects as an example, has grown dramatically in recent years. Glaeser and Scheinkman (2001) provide an overview of the state of the art in this literature. Two key insights appear in this literature which are particularly relevant to this paper.

The first, due to Manski (1993) is that a person generally chooses his or her social group. As a result, econometric treatments which treat the social group as exogenous produce biased parameter estimates. The second insight is that, in equilibrium, peer effects can dramatically increase the variability of group behavior. For example, Glaeser, Sacerdote, and Scheinkman (1996) apply an aggregate model with local interactions to the analysis of metropolitan crime rates. Their results support the hypothesis that social interactions create multiplier effects like those evaluated in this paper. Brock and Durlauf (2000) outline the aggregate implications of incorporating peer effects into a standard discrete choice model. The analysis in this paper applies a variation on Brock and Durlauf's model, along with enhancements I develop for use in applied work.

2 Data and methods

2.1 Data sources

The primary data source for this study is the Monitoring The Future (MTF) survey . This large national survey of high school seniors has been collected annually since 1976 by the Institute for Social Research. Prices are provided by the Tobacco Institute (1997) and represent the regional average price (including state and federal excise taxes) for a pack of cigarettes, in 1990 dollars, as of November 30 of the previous year. I also use two smaller surveys, the 1993 Teenage Attitudes and Practices Survey (TAPS) and the Youth Risk Behavior Survey (YRBS), to supplement the MTF on certain issues.

The 1993 TAPS survey features more detailed information than the MTF on factors such as parental behavior and discipline, school smoking policy, and exposure to various media influences. In addition to the greater detail, the TAPS survey is household-based, which provides a check on the possibility that the school-based MTF suffers from a composition bias due to the higher smoking rate of dropouts and the higher rate of dropout among black youth.

The YRBS is a school-based survey of 9th through 12th grade students which has been collected in 1990, 1991, 1993, 1995, and 1997. The YRBS has a more limited set of variables but does identify a respondent's state of residence. These state identifiers, unavailable in the other two surveys, enable me to investigate the possibility that the difference in smoking rates is attributable to a difference in prices, state-level regulations, or cigarette advertising across states combined with a difference in racial composition across states.

2.2 The model

The model used in this paper is a variation on the basic logit model in the spirit of Brock and Durlauf (2001). Agent i faces a binary choice, to be a smoker ($s_i = 1$) or nonsmoker ($s_i = 0$). The incremental utility from being a smoker takes the form:

$$u_i = \beta X_i + \gamma(f_i) + \epsilon_i \tag{1}$$

where X_i is a vector of variables including income, prices, and personal characteristics which may affect preferences or constraints, and f_i is the fraction of a person's friends who smoke. In this specification, γ is some differentiable parametric function to be estimated and β is some vector of parameters to be estimated. In estimation, no restriction is placed on the sign of $\gamma'(f_i)$. However, in discussing the behavior of equilibrium smoking rates, and deriving expressions for the social multiplier, I assume that peer effects are positive, i.e.:

$$\gamma'(f_i) \geq 0 \qquad \forall f_i \in [0,1] \tag{2}$$

The idiosyncratic term ϵ_i represents unobserved attributes. ϵ_i is independently and identically distributed across individuals with a logistic distribution and is independent of all observed attributes. Because one of the observed attributes is peer behavior, the independence assumption is quite restrictive. Section 2.4 discusses its implications. The probability that a person smokes is:

$$\Pr(s_i = 1 | X_i, f_i) = \Lambda(\beta X_i + \gamma(f_i)) \tag{3}$$

where:

$$\Lambda(X) = \frac{e^X}{1 + e^X}$$

The primary goal of this paper is to find an explanation - an economically justified model specification that, when used to form predictions from the data, reproduces the patterns in the actual time series. Whether a model specification has economic content is partly a matter of judgement. At the extreme end, a specification with race-specific year effects could reproduce the time series exactly but would be uninformative. In the spirit of Stigler and Becker (1977), I will avoid appealing to unexplained shifts in tastes and instead look for shocks to the choice environment.

Estimating the parameters of the model is slightly complicated by the fact that smoking is self-reported. As audit studies (Bauman et al. 1994, Wagenknecht et al., 1992) confirm, teenagers significantly understate their smoking in surveys. In these studies, youth are asked how often they smoke, and are later asked to provide a saliva sample which is tested for the presence of nicotine-related chemicals such as cotinine. The prevalence of underreporting in the TAPS and MTF data can be quantified by comparing the fraction of teens who say they smoke, and the fraction of their best friends that they report as smokers. In the TAPS data, 20% of respondents report that they smoke, but the average smoking rate among their best friends is 27%. Unless smokers have more friends, this implies that at least one-fourth of smokers in the TAPS survey self-report as nonsmokers.

Normally, the underreporting issue could be addressed simply by relabeling s_i as "reports smoking" rather than "smokes". However, the analysis of multiple equilibria and social multipliers will require that the measure of individual behavior (s_i) predicted by the model and the measure of peer behavior (f_i) refer to the same behavior. As a result, it is necessary to develop an explicit model of the relationship between being a smoker and self-reporting as a smoker.

Let r_i indicate whether person i reports smoking ($r_i = 1$) or not ($r_i = 0$), and let s_i indicate whether the person actually smokes. Assume that:

1. Respondents correctly report the smoking behavior of their friends.
2. The number of friends a teenager has is independent of whether he is a smoker or nonsmoker.
3. With probability $1 - R$, a smoker will falsely claim to be a nonsmoker. This probability is independent of all other observed characteristics.
4. Nonsmokers do not falsely claim to be smokers.

The variable s_i is not observed, but r_i and f_i are both observed. Equation (3) can be rewritten:

$$\begin{aligned} \Pr(r_i = 1|X_i, f_i) &= R \Pr(s_i = 1|X_i, f_i) \\ &= R\Lambda\left(\beta X_i + \gamma\left(f_i\right)\right) \end{aligned} \tag{4}$$

By the first two assumptions, $E(s_i) = E(f_i)$. By the last two assumptions, $E(r_i) = RE(s_i)$. Therefore:

$$R = \frac{E(r_i)}{E(f_i)} \tag{5}$$

I estimate the parameters of the model using a two-step procedure. First, I estimate R by substituting sample averages into equation (5). Then I substitute this estimate of R into equation (4) to estimate the remaining parameters.

2.3 Multiple equilibria and social multipliers

When a choice is subject to peer effects, the behavior of a group is properly understood as an equilibrium outcome of some game rather than as the aggregation of a set of independent individual decisions. As a result, group behavior may exhibit unusual dynamics - either rapid and unexplained shifts (multiple equilibria) or unexpectedly large reactions to aggregate shocks (social multiplier effects). This section formalizes these ideas, and describes how their empirical relevance can be assessed.

2.3.1 Definitions

First I define a peer group and equilibrium in a peer group.

Definition 2.1 (Peer group) *A peer group G is a probability distribution function for the vector of observable characteristics X.*

Definition 2.2 (Equilibrium) $s \in [0,1]$ *is an equilibrium smoking rate for a peer group G if it solves the equation:*

$$s = E_G\left(\Lambda\left(\beta X_i + \gamma\left(s\right)\right)\right) \tag{6}$$

Let $\bar{s}(G)$ be the set of all equilibrium smoking rates for group G.

I use G to refer both to the peer group itself and the distribution of observed characteristics within the group. While the equilibrium smoking rate is well-defined mathematically, it is useful to consider the economic situation I assume it describes. First, each member of the peer group is friends with every other member. In the language of social interactions models, interactions are global within the peer group. Second, each individual makes his smoking choice to maximize his utility taking the choices of others as given. Third, the peer group is large enough that several approximations – the realized smoking rate of the group, the realized smoking rate of the group not including the member making the smoking decision, and the expected value of these two quantities are all treated as the same quantity – are reasonable. This equilibrium concept is described in more detail in Brock and Durlauf (2000). Work in progress (Krauth 2001) analyzes the implications of assuming that peer groups are small.

Next I define a social multiplier. The idea of a social multiplier was formalized in Schelling (1978, p. 106), and reflects the idea that social interactions will amplify the effect of aggregate shocks on aggregate behavior. For example, consider the effects of a price decrease. Some teens may take up smoking as a direct result of the price decrease. Others may not change their behavior in response to the price decrease itself, but will take up smoking because the price decrease has increased the fraction of their friends that smoke. In this hypothetical case, the price decrease has both direct and indirect effects on the behavior of these teens.

Definition 2.3 (Social multiplier) *Let x be an arbitrary element of the vector X such that $\beta_x \neq 0$. Let G be a peer group such that $\bar{s}(G)$ is unique. The social multiplier for G is:*

$$m(G) \equiv \frac{\frac{d\bar{s}(G)}{dx}}{\left.\frac{\partial E_G(\Lambda(\beta X + \gamma(s)))}{\partial x}\right|_{s=\bar{s}(G)}}$$

Less formally, the social multiplier is simply the ratio of the total (direct and indirect) effect of an aggregate shock to the direct effect of the shock.

$$m(G) = \frac{\text{direct effect} + \text{indirect effect}}{\text{direct effect}}$$

2.3.2 Characterization of equilibria

In this section, I describe the characteristics of equilibria. In particular, I establish testable conditions for uniqueness and a formula for the social multiplier that can be calculated from parameter estimates. These results will be used in the empirical work of Section 3.

Existence of equilibrium follows from the observation that the right side of equation (6) is a continuous mapping from the unit interval to itself.

Proposition 2.1 (Existence) *$\bar{s}(G)$ is nonempty for all G.*

Given a parameter vector (β, γ) and a peer group G, one can find if there are multiple equilibria through a simple algorithm: starting with either one or zero, iterate on equation (6) until convergence. If the iteration converges to the same point starting from both one and zero, equilibrium is unique;[3] otherwise it is not. It is also possible to determine, for a given parameter vector, whether there exists any G for which there are multiple equilibria. The following sets of propositions outline the conditions.

Proposition 2.2 (Uniqueness vs. multiplicity) *$\bar{s}(G)$ is unique for any G if:*

$$\gamma'(s)\,\Lambda'\left(\Lambda^{-1}(s)\right) < 1$$

or (equivalently)

$$\gamma'(s)\,s(1-s) < 1$$

for all $s \in [0,1]$. If this condition is not met, then $\bar{s}(G)$ is nonunique for some G.

Corollary 2.3 (Uniqueness in linear case) *Let $\gamma(s) = \gamma s$. Then $\bar{s}(G)$ is unique for any G if $\gamma < 4$. Otherwise, $\bar{s}(G)$ is nonunique for some G.*

Corollary 2.4 (Uniqueness in quadratic case) *Let $\gamma(s) = \gamma_1 s + \gamma_2 s^2$. Then $\bar{s}(G)$ is unique for any G if:*

$$\gamma_2 < \frac{1 - \gamma_1 s(1-s)}{2s^2(1-s)}$$

for all $s \in [0,1]$. Otherwise, $\bar{s}(G)$ is nonunique for some G.

[3] Provided that $\gamma(.)$ is nondecreasing, which holds in all of the empirical results here.

Proposition 2.2 and Corollaries 2.3 and 2.4 define regions of the parameter space in which equilibrium must be unique. Outside of that region, multiple equilibria are possible. Figure 2 depicts the multiple equilibria region for the quadratic functional form used in this paper.

The social multiplier can be found as a function of model parameters and group characteristics:

Proposition 2.5 (Social multiplier) *The social multiplier for a group G is:*

$$m(G) = \left. \frac{1}{1 - \gamma'(s) E_G(\Lambda'(\beta X + \gamma(s)))} \right|_{s=\bar{s}(G)}$$

The formula in Proposition 2.5 can be considered an empirically tractable special case of the more general calculations in Cooper and John (1988), and in Glaeser and Scheinkman (2000)

Because the data sets used in this study (and in most studies of peer effects) do not include complete information on the composition of a person's peer group, it is necessary to make assumptions about that composition. In this paper, I assume that peer groups are homogeneous in terms of observed characteristics; i.e., $X_i = X_j$ for all i, j in the group. In this case, Proposition 2.5 can be simplified:

Corollary 2.6 *If all members of G have the same observable characteristics, then:*

$$m(G) = \left. \frac{1}{1 - \gamma'(s)\, s(1-s)} \right|_{s=\bar{s}(G)}$$

Under this assumption, each observation has its own estimated social multiplier, so I report the median social multiplier. An alternative is to assume that each individual chooses friends randomly from the population as a whole. In this case, the social multiplier is constant across the population. The resulting estimates (not reported) are similar to those calculated under the assumption of homogeneous peer groups.

2.3.3 Intuition

To develop intuition for the propositions in this section, consider the special case that $\gamma(s) = \gamma s$. The right side of equation (6) can be thought of as a response function - the smoking probability of a representative group member as a function of the smoking rate of the rest of the group. Figure 3 depicts this response function for different levels of γ. An equilibrium is simply any point at which this function crosses the 45-degree line. As γ increases (from 3 in the first graph to 8 in the second), the response function gets steeper and multiple equilibria become possible. Looking at the graphs, it is easy to see that equilibrium is unique if the slope of the response function is always less than one – because once it has crossed the 45-degree line, which has a slope of one, it must have a higher slope to cross it again. This result is formalized in Proposition 2.2 and its corollaries.

Figure 4 depicts the social multiplier graphically. A small aggregate shock will have the effect of shifting the response function upwards or downwards. The size of the vertical shift corresponds to the direct effect, that is, the increase in smoking probabilities of a representative group member keeping peer behavior constant. However, peer behavior is not constant and so there is an additional indirect effect on the group's equilibrium smoking rate. The social multiplier, which is the ratio of total effect to direct effect, is equal to $\frac{1}{1-b}$, where b is the slope of the response function. This intuition is formalized in Proposition 2.5, and is conceptually no different from the "Keynesian cross" seen in first-year undergraduate macro.

2.4 Endogeneity of peer group

One potential problem with the estimation procedure is that peer behavior is treated as an exogenous variable for the purposes of estimation. In other words, the disturbance term ϵ_i is independent across peer group members and peer groups are large. Unfortunately, neither of these assumptions is very reasonable for this case. Unobserved variables are likely to be highly correlated among members of a peer group as a result of self-selection, and the relevant social group which affects decisions about smoking is likely to be small. Manski (1993), among others, analyzes the self-selection problem in detail, while Krauth (2001) discusses the implications of small-group interactions. In both cases, the estimated impact of peer behavior is likely to overstate the true impact.

Ideally this endogeneity problem could be corrected through use of an instrumental variable. Unfortunately, an acceptable instrument would have to be a variable which affects the smoking rate of a person's friends, but not his or her own smoking rate – a criterion that is simply not met by any variable in either the TAPS or MTF data. As a result of these issues, the marginal impact of peer behavior implied by the coefficient estimates should should be considered an upper bound on the true marginal impact. The estimated social multiplier, which is an increasing function of the marginal impact, should also be considered an upper bound. While this is somewhat dissatisfying, it may be preferable to a questionable instrument or structural model. In addition, one of the main results in this paper is that peer influence is not strong enough to imply multiple equilibria. Clearly, this negative result is not weakened by the knowledge that the estimated impact of peer influence is likely to be overstated.

2.5 Measuring peer behavior

While the empirical analysis of multiple equilibria and social multipliers requires quantitative measures of peer behavior, the MTF questions on peer smoking are qualitative in nature. This section describes the procedure used to impute quantitative measures of peer behavior from the qualitative responses. The procedure exploits the overlap between the TAPS and MTF populations combined with the quantitative data on peer behavior in the TAPS survey.

The following model will provide the mapping. Each student has a large number of friends, some fraction f_i^{MTF} which smoke. The MTF survey asks respondents "How many of your friends would you estimate smoke cigarettes?" with potential responses of "none", "a few", "some", "most", and "all". Assume that f_i^{MTF} takes on one of five values, each of which corresponds to an answer on the MTF questionnaire.[4]

[4]Assuming a five-point distribution here is strictly a matter of convenience. I could instead assume that each student has the same range in his mind for these five categories and define f_i^{MTF} as the expected fraction of i's friends who smoke conditional on their answer to the MTF question.

Answer	f_i^{MTF}
None	0
A few	F
Some	S
Most	M
All	1

As these categories have a natural ordering, I also assume that:

$$0 \leq F \leq S \leq M \leq 1 \tag{7}$$

The TAPS survey asks respondents exactly how many of their four best friends smoke. Assume that a student's four best friends are selected randomly with replacement from his/her group of friends. As a result, the number of four best friends who smoke is a random variable with a binomial distribution.

If a researcher were to ask the MTF question and the TAPS question to members of the same population, the probability of a randomly selected respondent saying that X of his four best friends smoke would be:

$$\begin{aligned} \Pr(f_i^{TAPS} = X) &= \Pr(f_i^{MTF} = 0)B_{4,0}(X) \\ &+ \Pr(f_i^{MTF} = F)B_{4,F}(X) \\ &+ \Pr(f_i^{MTF} = S)B_{4,S}(X) \\ &+ \Pr(f_i^{MTF} = M)B_{4,M}(X) \\ &+ \Pr(f_i^{MTF} = 1)B_{4,1}(X) \end{aligned} \tag{8}$$

where f_i^{TAPS} is person i's response to the TAPS question, f_i^{MTF} is his response to the MTF question, and $B_{n,p}(X)$ is the probability of drawing X from the binomial distribution with parameters n and p.

The MTF and TAPS describe different populations. However, it is possible to take subsamples of each data set that are both random samples of the population of 1993 respondents who were in the 12th grade. Given the marginal distributions of f_i^{TAPS} (estimated from the TAPS data) and f_i^{MTF} (estimated from the MTF data) in this population, the parameters (F, S, M) maximize the following log likelihood:

$$\ln L = \sum_{j=0}^{4} TAPS_j \ln \left[\sum_{k \in \{0,F,S,M,1\}} MTF_k B_{4,k}(j) \right] \tag{9}$$

subject to the constraint (7), where $TAPS_j$ is the number of TAPS respondents that said j of their four best friends smoke, and MTF_k is the fraction of MTF respondents that answered k to the question "how many of your friends smoke?". The results are:

$$\begin{aligned} \hat{F} &= 13\% \qquad (1.01\%) \\ \hat{S} &= 41\% \qquad (1.55\%) \\ \hat{M} &= 90\% \qquad (1.15\%) \end{aligned}$$

Standard errors of the estimates are reported in parentheses. I use these results to recode the MTF data, coding anyone who says "a few" of his friends smoke as having a friends' smoking rate of 13%, anyone who says "some" as 41%, and so on. Figure 5 shows the frequency distribution of f_i in the TAPS data, along with the frequency distribution implied by substituting the maximum likelihood estimates into equation (8). As the figure shows, this simple specification fits the data quite well.

3 Results

3.1 Results without peer effects or differential price response

Table 1 shows logit model estimates for a baseline specification using explanatory variables including prices, basic demographic and regional characteristics, parental education, church attendance, and disposable income. The baseline specification assumes no peer effects and all coefficients are constrained to be identical for black and white teenagers. The first column shows results for the MTF survey, while the second shows results for the TAPS survey. Price is omitted from the TAPS specification because the TAPS has only one year of data, so price is perfectly collinear with region.

Most of the coefficient estimates agree in sign with both the existing literature and with expectations. Smoking increases with disposable income and age, and decreases with prices and church attendance. Surprisingly, higher parental education is associated with a slightly higher smoking probability, even though it is well known that smoking rates in adults are decreasing in education. One possible explanation is that parental education acts as a proxy for parental income. The MTF estimates have quite small standard errors due to the large sample size. Table 2 includes some additional TAPS variables which are not available in the MTF survey. In these results, smoking is increasing in parental smoking, teacher smoking, and family income, and decreasing in sports participation, discussion of the risks of smoking in class, and parental education. Interestingly, students who report that they are exposed to antismoking messages in television and radio are more likely to smoke - this may be simply a matter of these messages being more memorable to smokers.

Do these models explain the difference in smoking rates? To answer this question, I use the parameter estimates to calculate predicted smoking probabilities for each observation, then take group averages to get the "Predicted smoking rate" items in Tables 1 and 2. As the tables show, the parameter estimates imply very little difference in smoking rates between white and black teens. The predicted smoking rate can also be calculated on an annual basis. Figure 6 shows the time series of smoking rates for black and white teenagers predicted in the MTF survey at the estimated parameter values. The model predicts very little difference in smoking rate between black and white teenagers. The reason for this is simple: few of the variables included changed substantially for black teens while remaining the same for white teens, and those that did (parental education, for example) simply don't have a quantitatively important impact on smoking probabilities in the estimated model.

What are the implications of these results? None of the variables in these regressions (urban/suburban/rural, disposable income, parental income, parental education, parental smoking, exposure to antismoking messages, teacher behavior, alternative activities like religious participatation or sports) is particularly useful in explaining the divergence in teen smoking rates.

3.2 Peer behavior

Next we consider peer effects and their implications. Table 3 shows parameter estimates for the baseline model with a quadratic peer effect added. The peer effect is large and statistically significant.

To see if peer behavior matters enough to produce multiple equilibria, I plot the parameter estimates, with 95 percent joint confidence ellipses, in Figure 7. If the parameters fall in the shaded area, there are multiple equilibria for some (not necessarily all) groups. If the parameter values fall in the white area, then equilibrium is unique for all groups. The MTF results indicate that equilibrium is unique, while the TAPS point estimates lie in the multiple equilibria range. However, the TAPS estimate is close to the multiple equilibria frontier; near the frontier, most groups will have unique equilibria. This can be quantified in the following way. Assume that each individual is in a group which has the exact same observable characteristics as himself. It is then possible to determine whether such a group has a unique equilibrium smoking rate. At the TAPS point estimates, only 3.24% of respondents would be in a group with multiple equilibria. In addition, the confidence interval is relatively large, and, as discussed in Section 2.4, the point estimates should be considered upper bounds on the true peer effect. As a result, Figure 7 implies that the equilibrium-switching story is simply not supported by the data.

Having ruled out the multiple equilibria explanation, I estimate the social multiplier. The social multiplier for the median peer group is reported in Table 3. The estimated social multiplier is 2.5, indicating that the indirect effect of an aggregate shock is larger than the direct effect. An aggregate shock which would directly raise a typical teen's smoking probability by 1 percent would raise the smoking rate of a peer group (or collection of peer groups) by 2.5 percent. As a result, given an aggregate shock that explains about half of the difference in smoking rates between white and black teens, the social multiplier process would explain the other half. The next section suggests such an aggregate shock.

3.3 Differential price response

In this section I consider the effect of relaxing the assumption that parameters are race-invariant. This does not imply an assumption that preferences differ systematically by race, but rather an acknowledgement that important missing variables may be correlated with race.

First, I include race as an explanatory variable in the baseline regression. In the interests of space, the parameter estimates are not reported; the time series of predicted smoking rates is shown in Figure 8. As the figure shows, this modification shifts the time series for black teens down but doesn't produce any time series trends like those in Figure 1. This should not be surprising, but it emphasizes the point that any explanation must do more than find differences between black and white teenagers, or posit the existence of unobserved differences. An acceptable explanation must have some time series dimension to it.

The next possible explanation is that price response varies between black and white teenagers. To see how this might matter, Figure 9 shows the time series of cigarette prices in 1990 dollars. As the figure shows, prices rose sharply during the 1980's and stopped rising in the early 1990's. Standard preferences would imply a sharp fall in smoking rates during the 1980's as a result, exactly the pattern for black teenagers. Chaloupka and Pacula (1998) suggest that racial differences in price elasticity combined with rising prices during the 1980's could explain the time series patterns in youth smoking. If black teens are more responsive to price, then one would expect them to experience a steep drop in smoking during the 1980's while white teens experience a smaller drop.

To evaluate this explanation, I estimate a model in which price response varies by race. The results are reported in Table 4, both with and without peer behavior. As the table shows, black teenagers seem to have a much larger price response in both cases. This finding of differential price response in black and white teens is quite consistent with the existing literature. Chaloupka and Pacula (1998) and Gruber and Zinman (2001), using data with much more detail on prices, both find substantially higher price elasticities for black teens. One potential issue with the analysis here is that the MTF data do not include state identifiers, so prices are regional averages. In contrast, these other authors use data sets with state identifiers. Chaloupka and Pacula use a restricted use version of the 1992-1994 MTF surveys that have state identifiers, while Gruber and Zinman use the less detailed YRBS survey that I use in Section 3.4.2. While each of these three data sets have their own flaws, the finding of substantially higher price responses by black teens appears fairly robust.

The finding of a higher price response by black teens is not, however, the end of the story. Few social scientists are content with simply positing an unexplained difference in preferences between two social groups. As a result it is necessary to look for differences in the choice environment which, once accounted for, reduce the unexplained difference in elasticities.

First, we consider the role of peers. Note that the difference in price response is lower once peer behavior is included as an explanatory variable. An intuitive explanation for this is that the coefficient on prices already includes the multiplier effect when peer behavior is omitted, but does not include the effect when peer behavior is included.[5] As Table 4 shows, accounting for the social multiplier effect reduces the difference in elasticities to be explained by nearly half. We can also look at the interaction of peer behavior and prices graphically. Figure 10 shows the time series of predicted smoking rates for the parameter estimates with peer influence. The figure shows a large drop during the 1980's for black teens and a small drop for white teens, exactly the pattern which appears in the data. Note that some of the drop is due to the price change and some of it is due to the social multiplier effect. Figure 11 shows predicted smoking rates from the model with peer behavior, but holding peer behavior constant. As the figure shows, the social multiplier effect from peer behavior plays an important quantitative role in this story, even though the price change is the driving force.

Next, I investigate possible explanations for the difference in price response by interacting price with disposable income, parental education, and whether the respondent lives in an urban environment. The results are shown in Table 5. As the table shows, disposable income seems to have no economically significant relationship to price elasticity. In addition, there is almost no difference in the disposable income reported by black and white teens in the MTF. While this

[5]It should be emphasized that this intuition is strictly informal as the coefficients of misspecified logit models are complicated functions of the structural parameters.

might run counter to intuition, disposable income in the MTF is composed of allowance and labor income. Black teens in the survey make slightly lower labor income, but receive larger allowances, a result that is confirmed in a detailed analysis of allowances in the NLSY97 survey by Pabilonia (1999). Differences in current disposable income thus cannot explain the difference in price response.

What other explanations could explain such a large difference in price response? One possibility, suggested by the rational addiction approach, is that smoking responds to expectations of future income and prices. In this case, the lower lifetime income of black adults, combined with a strong association between current prices and expected future prices, could explain the higher price elasticity of black teens. Another possibility is that the cost of smoking to teens is not exclusively a matter of price. Because of laws forbidding the sale of tobacco to minors, there may be differences in the non-monetary cost of smoking for white and black teens. As a result, price elasticities may differ without unobserved preference heterogeneity.

3.4 Alternative explanations

Several potentially appealing explanations are best addressed separately from the basic approach of this paper. In this section, I address those explanations using other forms of evidence.

3.4.1 Underreporting bias

As discussed earlier, smoking behavior is self-reported in all three surveys, leading to a potential source of bias in the results. If black teens have higher underreporting rates, the difference in self-reported smoking rates will be overstated. It could even be the case that black and white teens have the same smoking rates, but differences in underreporting create a spurious difference in the self-reported rates. This section investigates that question.

The audit studies described in Section 2.2 find that black teenagers underreport their smoking more often than white teens. However, it is unlikely that underreporting bias should play a large role in explaining the divergence in smoking rates. First, the degree of differential underreporting needed to generate the observed gap in self-reported smoking rates is quite large. Even if no white smokers falsely claim to be nonsmokers, at least two-thirds of black smokers would have to falsely report not smoking in order to explain the youth smoking differential in 1993. If some white smokers falsely claim to be nonsmokers, that fraction increases. Second, using differential underreporting to explain the widening gap from 1980 to 1990 would require corresponding large shifts in the underreporting rate.

As in Section 2.2, I use reports on friends' behavior to estimate underreporting rates. In the TAPS data, approximately 16.2% of the friends of black teens smoke, compared with approximately 29.5% of the friends of white teens. 7.7% of black teens report current smoking while 22.5% of white teens report current smoking. If social circles are racially homogeneous, that implies that 29.5% of white teens and 16.2% of black teens smoke, and that 24% of white smokers and 52% of black smokers falsely claim to be nonsmokers. In other words, black teens do have higher underreporting rates. However, the implied gap in smoking rates is still quite large. A similar exercise can be used to detect time trends in underreporting in the MTF survey. Figure 13 shows the time series of underreporting rates by race (assuming social segregation). While the underreporting rate for black teens increases somewhat during the 1980's, it does not increase dramatically.

If the assumption of social segregation is relaxed, differential underreporting plays a smaller role. The friends' smoking rate can be interpreted as a weighted average of the smoking rates of black and white teenagers, and the calculations above reflecting an assumption that the weights are zero and one. With the (unobserved) correct weighting, the estimated "true" smoking rate of black teens will fall and the estimated true smoking rate of white teens will rise. This will also imply a fall in the underreporting rate of black teens and a rise in the underreporting rate of white teens. In the MTF time series, this implies there will be an increase over time in measured underreporting for black teens and a decrease over time for white teens if social segregation is falling. In other words, the numbers calculated under the assumption of social segregation will tend to overstate the effect of differential underreporting. Even these overstated results still support the qualitative pattern of smoking rates in Figure 1, implying that differences in underreporting do not explain the pattern.

3.4.2 Differences in prices and regulations

Cigarette prices vary widely across states due to differences in excise taxes. In 1998, for example, a pack of cigarettes sold for approximately \$1.55 in Kentucky and \$2.73 in Washington, a difference of 76%. In addition, laws regulating where a person can smoke, where cigarettes can be sold, etc., also vary widely across states. The MTF data do not provide state identifiers, so regional average prices are used to measure the prices faced by each respondent. It is possible that the distribution of black teens across states implies that they face higher prices, and that these higher prices are not captured by regional averages.

I use the YRBS survey with linked state-specific and year-specific prices to evaluate this explanation. States also differ in the regulations they attach to smoking, for example, whether vending machines are allowed. As is commonly pointed out in the literature, these policies are highly correlated with both taxes and with attitudes towards smoking - the states with the lowest taxes and fewest regulations on cigarettes tend to be tobacco-growing Southern states like Kentucky and North Carolina. If regulations and attitudes are omitted variables, as they are here, the effects of price are likely to be overstated. However, the purpose of this exercise is not to estimate price elasticities, but to account for how much of the difference in smoking rates is due to state of residence. The sale price of tobacco can thus be considered a summary measure of the total cost (counting regulations) of smoking. Table 6 reports the results for a logit model in which the price is included as a regressor. Figure 14 shows the time series of predicted smoking rates for the YRBS over the period 1990-1997. As the figure shows, differences in prices do not explain the lower smoking rate of black teens during this period.

There is an alternative and more general approach to this same question. I estimate a linear probability model of the form:

$$\Pr(s_i = 1) = \alpha b_i + \beta X_i \tag{10}$$

where X_i is a vector of state-year dummy variables. In other words, this model allows for a year-specific state effect, which will include all differences in prices, regulations, attitudes, and advertising across states. The coefficient α can be interpreted as the average difference in black-white teen smoking rates within a state in a given year over the sample period. The resulting estimate of α is -0.17 with a standard error of 0.0043, implying that the average difference in smoking rates within a state is about 17%. Clearly, the distribution of black teens across states does nothing to explain the difference in their smoking rates.

3.4.3 Advertising

One possibility is that reduced exposure to cigarette advertising could explain the decrease in the black teen smoking rate. Unfortunately, detailed historical data on advertising of the sort that could address this question are not available. The analysis in Section 3.4.2 indicates that differences in advertising rates across states could not possibly explain the difference, since within-state differences in smoking rates are comparable to national differences. However, it is possible to assess the plausibility of that hypothesis. By agreement with the plaintiffs in the various late 1990's lawsuits against the tobacco industry, the Tobbacco Institute maintains a searchable web archive (`http://www.tobaccoinstitute.com`) of all its internal documents that have been subpoenaed by the plaintiffs. This archive includes memos, internal research reports, and hundreds of newspaper and magazine clippings related to the tobacco industry. A search for articles on "black" or "African-American" and "advertising" yielded numerous articles from the period 1985-1995 specifically claiming that the tobacco industry advertised more to African Americans than to whites. These include a 1986 CBS Evening News report titled "Black community targeted by cigarette companies", a 1990 editorial in the Journal of the National Medical Association, a 1986 Washington Post editorial, an Associated Press story in 1989, and stories in the Houston Post and Miami Herald in 1990, the Fort Lauderdale Sun-Sentinel in 1991, and the St. Louis Post-Dispatch in 1995. No articles were found suggesting that there had been a decrease in advertising to African-Americans. While this is not convincing evidence that black teens really were exposed to more cigarette advertising than white teens, it certainly seems likely that a decline in advertising in the black community large enough to cause a 50% decline in the teen smoking rate would have been noticed by someone. This suggests that a large decline in advertising is unlikely. Indeed, some analysts (McIntosh 1995) have suggested that the flood of targeted advertising, including an abortive attempt by R.J. Reynolds to introduce a new cigarette brand ("Uptown") openly aimed at black smokers, led to a backlash which created an anti-smoking social norm among young African-Americans. While this is an intriguing hypothesis, the beginning of the decline in cigarette smoking among black teens predates this controversy by several years.

4 Conclusion

The dramatic fall in smoking rates among black teens in the 1980's can be attributed to three factors - a rise in the price of cigarettes, a higher price elasticity among black teens, and the amplifying effects of peer influence. Most of the obvious explanations for the differences in behavior between white and black teens – prices, regulations, parental behavior, exposure to anti-smoking messages, and disposable income – can be dismissed. Other explanations, including exposure to advertising and differential underreporting, cannot be completely dismissed as factors but are unlikely to be the driving force. Finally, the possibility that the fall could be explained through equilibrium-switching was found inconsistent with the data.

In addition to these substantive results, the paper develops several tools for the analysis of social interactions - a method for imputing quantitative measures of peer behavior from qualitative data, a method for determining whether parameter estimates are consistent with multiple equilibria, and a method to estimate social multipliers from the parameter estimates of a standard model of peer effects.

However, many open questions remain. The paper has also found that social interaction effects may play an important amplifying role, but the estimates of peer influence suffer from a difficult-to-quantify degree of sorting bias. Further research is thus needed to find more credible estimates of peer effects in teen smoking. The difference in price response between black and white teenagers has significant implications for public health policy. More research is needed to determine the cause of this difference, and whether the high price elasticity of black teens is due to policy or parenting choices which can be duplicated for white teens, or whether it is due to idiosyncratic factors which cannot be duplicated.

References

Bauman, Karl E. and Susan E. Ennett, "Tobacco use by black and white adolescents: the validity of self-reports," *American Journal of Public Health*, 1994, *84*, 394–398.

Becker, Gary S. and Kevin M. Murphy, "A theory of rational addiction," *Journal of Political Economy*, 1988, *96*, 675–700.

Brock, William A. and Steven N. Durlauf, "Discrete Choice with Social Interactions," *Review of Economic Studies*, 2001. forthcoming.

Chaloupka, Frank J. and Rosalie L. Pacula, "An examination of gender and race differences in youth tobacco smoking responsiveness to price and tobacco control policies," Working Paper 6541, NBER 1998.

Cooper, Russell W. and Andrew S. John, "Coordinating Coordination Failures in Keynesian Models," *Quarterly Journal of Economics*, 1988, *103*, 441–63.

Faulkner, Dorothy L. and Robert K. Merritt, "Race and Cigarette Smoking Among United States Adolescents: The Role of Lifestyle Behaviors and Demographic Factors," *Pediatrics*, February 1998, *101* (2).

__ , Luis G. Escobedo, Bao-Ping Zhu, Jeffrey H. Chrismon, and Robert K. Merritt, "Race and the incidence of cigarette smoking among adolescents in the United States," *Journal of the National Cancer Institute*, August 21, 1996, *88* (16), 1158–1160.

Glaeser, Edward L. and José A. Scheinkman, "Non-Market Interactions," Working Paper 8053, National Bureau of Economic Research 2000.

__ and __ , "Measuring Social Interactions," in Steven N. Durlauf and H. Peyton Young, eds., *Social Dynamics*, Brookings Institution, 2001, pp. 84–131.

__ , Bruce Sacerdote, and José A. Scheinkman, "Crime and Social Interactions," *Quarterly Journal of Economics*, 1996, *111* (2), 507–548.

Gruber, Jonathan and Jonathan Zinman, "Youth smoking in the U.S.: Evidence and implications," in Jonathan Gruber, ed., *Risky Behavior Among Youths: An Economic Analysis*, University of Chicago Press, 2001.

Johnston, Lloyd D., Jerald G. Bachman, Patrick M. O'Malley, John E. Schulenberg, and John Wallace, *Monitoring the Future Study, 1975-1998: United States*, Ann Arbor, MI: Institute for Social Research, University of Michigan, 1998.

Krauth, Brian V., "Small neighborhoods," Working Paper, Simon Fraser University 2001.

Manski, Charles F., "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 1993, *60*, 531–542.

McIntosh, Hugh, "Black teens not smoking in great numbers," *Journal of the National Cancer Institute*, 1995, *87*, 4.

Nelson, David E., Gary A. Giovino, Donald R. Shopland, Paul D.

Mowery, Sherry L. Mills, and Michael P. Eriksen, "Trends in cigarette smoking among US adolescents, 1974-1991," *American Journal of Public Health*, January 1995, *85* (1), 34–40.

Novotny, Thomas E., "Smoking among black and white youth: Differences that matter," *Annals of Epidemiology*, November 1996, *6* (6), 474–475.

Pabilonia, Sabrina W., "The role of the family in determining youth employment," Working Paper, University of Washington 1999.

Schelling, Thomas, *Micromotives and Macrobehavior*, W.W. Norton, 1978.

Stigler, George J. and Gary S. Becker, "De gustibus non est disputandum," *Journal of Political Economy*, 1977, *67*, 76–90.

The Tobacco Institute, *The Tax Burden on Tobacco: 1997*, Washington: The Tobacco Institute, 1997.

U.S. Department of Health and Human Services, National Center for Health Statistics, *Teenage Attitudes and Practices Survey, 1993: United States*, Ann Arbor MI: Inter-university Consortium for Political and Social Research, 1995. Computer file.

___ , *Youth Risk Behavior Survey, 1990-1997: United States*, Washington: Centers for Disease Control and Prevention, 1997. Computer file.

van Roosmalen, Erica H. and Susan A. McDaniel, "Peer Group Influence as a Factor in Smoking Behavior of Adolescents," *Adolescence*, 1989, *24*, 801.

___ **and** ___ , "Adolescent Smoking Intentions: Gender Differences in Peer Context," *Adolescence*, 1992, *27*, 87.

Wagenknecht, K.E., G.L. Burke, L.L. Perkins, N.J. Haley, and G.D. Friedman, "Misclassification of smoking status in the CARDIA study: a comparison of self-report with serum cotinine levels," *American Journal of Public Health*, 1992, *82*, 33–36.

Wang, Min Qi, Eugene C. Fitzhugh, R. Carl Westerfield, and James M. Eddy, "Family and Peer Influences on Smoking Behavior Among American Adolescents: An Age Trend," *Journal of Adolescent Health*, March 1995, *16*, 200–203.

A Figures

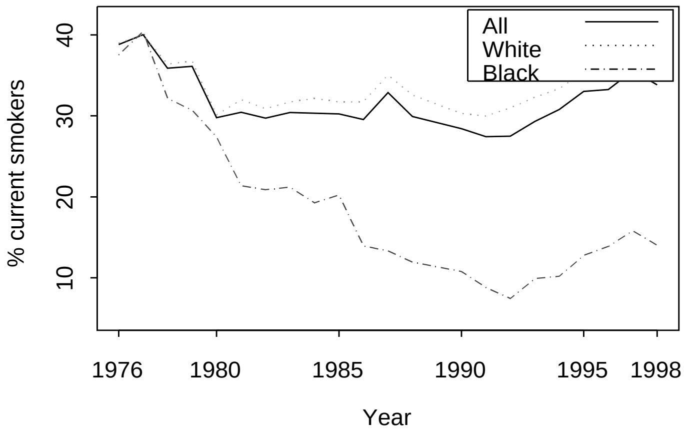

Figure 1: Smoking rates by race.

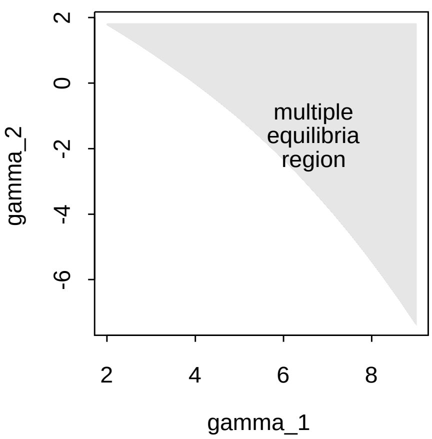

Figure 2: Critical range for parameter values in the quadratic case. If parameter values fall in the unshaded region, equilibrium is always unique. In the shaded region, equilibrium will not be unique for some distribution of characteristics.

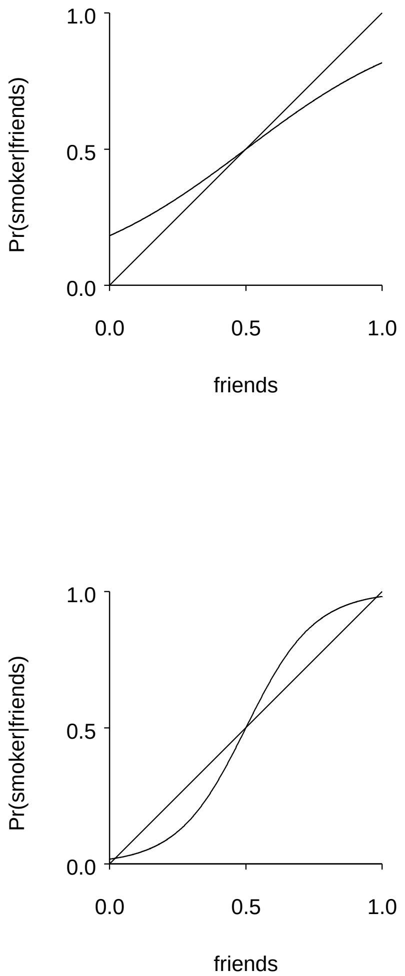

Figure 3: Finding a group equilibrium. Each graph depicts the "response function" (the probability of a typical group member smoking as a function of the smoking rate of the rest of the group) implied by a particular value of the peer effect γ. An equilibrium is simply a point where these two things are equal. In the first graph ($\gamma = 3$) equilibrium is unique, while the second graph ($\gamma = 8$) shows multiple equilibria.

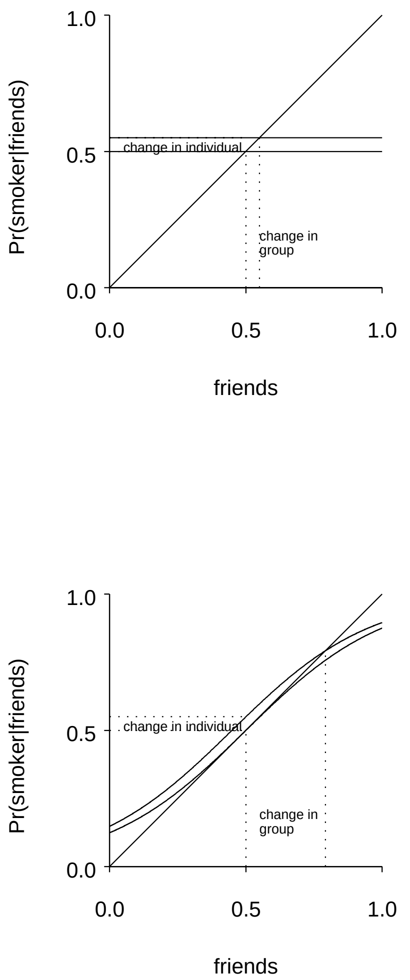

Figure 4: The social multiplier. As the response function shifts upwards, the equilibrium shifts as well. When the response function is relatively flat, as in the first graph, the change in group smoking rate for a small shock is small. When the response function is steeper, as in the second graph, the change in group smoking rate is larger. Note that the size of the upwards shift indicates the marginal effect keeping peer behavior constant.

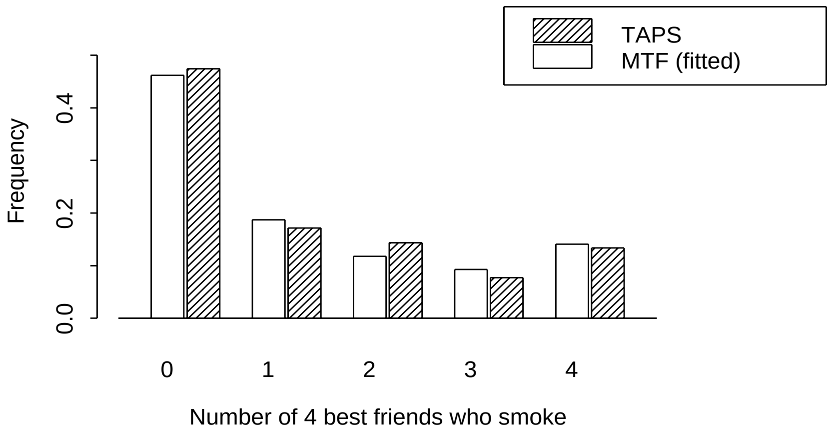

Figure 5: Distribution of “How many of your four best friends smoke?” in TAPS data, and implied distribution in MTF data for the mapping derived in Section 2.5

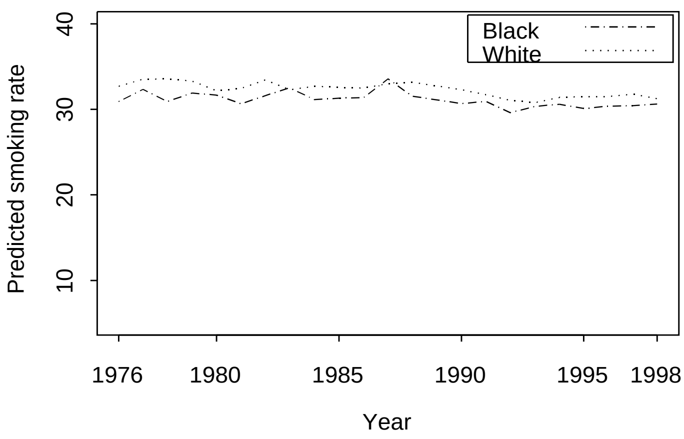

Figure 6: Predicted smoking rates for baseline model.

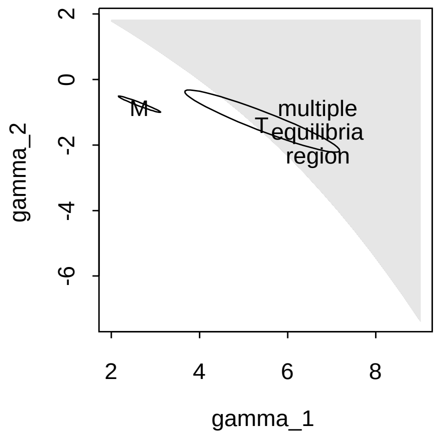

Figure 7: Parameter estimates for peer effect with 95 percent joint confidence ellipses. Parameters in unshaded range imply equilibrium is unique. "M" indicates MTF point estimate, "T" indicates TAPS point estimate.

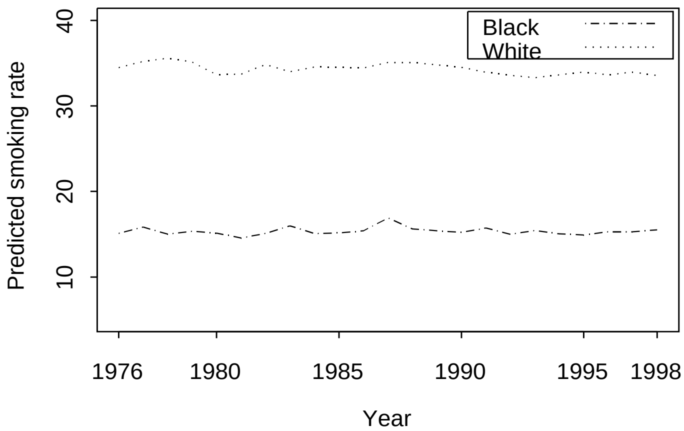

Figure 8: Predicted smoking rates for baseline model with race included as an explanatory variable.

Figure 9: Average price of cigarettes, 1990 dollars.

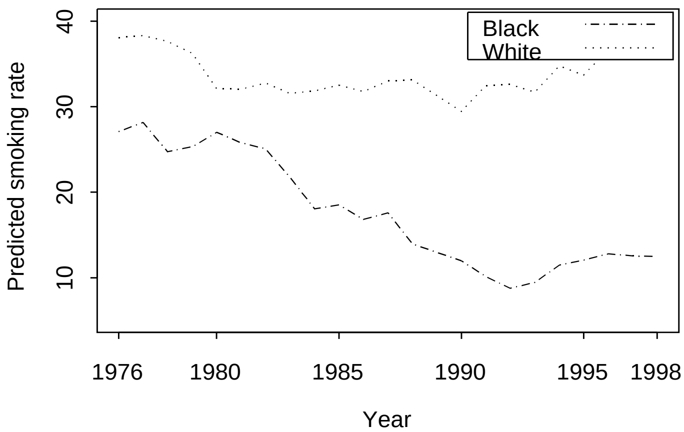

Figure 10: Predicted smoking rates for model with price response allowed to vary by race.

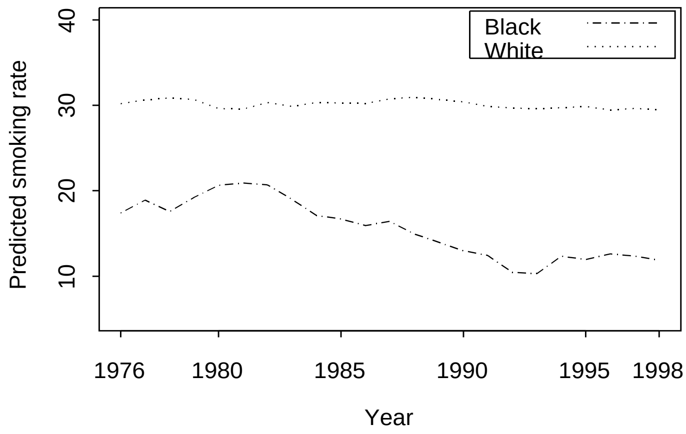

Figure 11: Predicted smoking rates for model with price response allowed to vary by race and peer behavior held constant

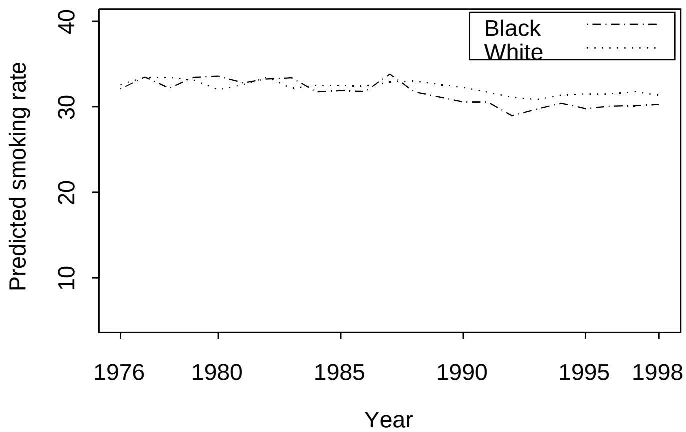

Figure 12: Predicted smoking rates for baseline model with price response interacted with income, parental education, and urban indicator.

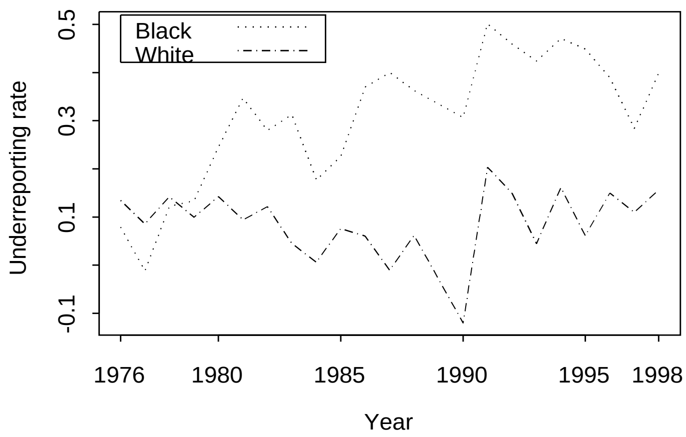

Figure 13: Time series of underreporting rates, estimated under assumption of social segregation.

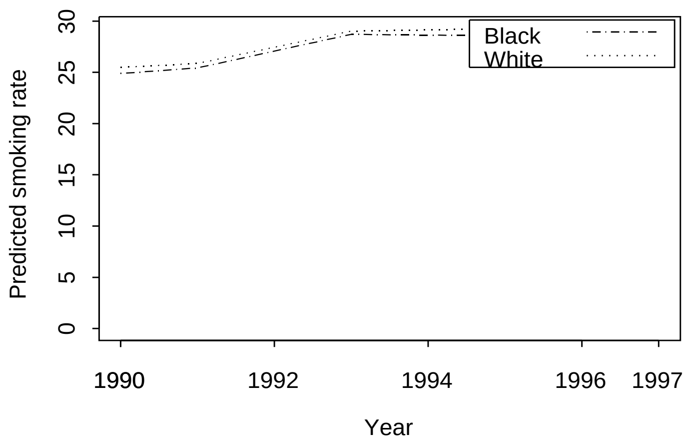

Figure 14: Predicted smoking rates in YRBS data.

B Tables

Variable	MTF Baseline	TAPS Baseline	Variable	MTF Baseline	TAPS Baseline
Intercept	-0.202	-3.823	Male	-0.145	-0.029
	(0.050)	(0.501)		(0.013)	(0.062)
Age	0.048	0.211	Parent HS	0.073	-0.160
	(0.014)	(0.029)		(0.025)	(0.097)
Parent College	0.043	-0.033	Church	-0.741	-0.870
	(0.014)	(0.073)		(0.016)	(0.063)
Urban	-0.306	-0.341	Suburban	-0.034	0.041
	(0.020)	(0.090)		(0.014)	(0.075)
Disp. Inc.	0.004	0.002	Midwest	0.020	0.082
	(1.26e-4)	(4.95e-4)		(0.019)	(0.087)
South	-0.207	-0.128	West	-0.447	-0.211
	(0.018)	(0.089)		(0.021)	(0.098)
Price	-0.185				
	(0.028)				
Predicted (black)	0.3077	0.1731			
Predicted (white)	0.3193	0.2044			
# of Observations	120046	8241			

Table 1: Baseline results.

Variable		Variable	
Intercept	-4.106	Male	0.084
	(0.764)		(0.064)
Age	0.189	Parent HS	-0.169
	(0.030)		(0.106)
Parent College	0.089	Church	-0.758
	(0.078)		(0.064)
Urban	-0.365	Suburban	0.065
	(0.093)		(0.077)
Disp. Inc.	0.003	Midwest	0.114
	(5.17e-4)		(0.089)
South	-0.154	West	-0.121
	(0.091)		(0.101)
Hispanic	-0.442	Log Family Income	0.050
	(0.142)		(0.057)
Parental Smoking	0.407	Media Exposure	0.533
	(0.064)		(0.067)
No Smoking Rule	-0.042	Teachers Smoking	0.298
	(0.085)		(0.066)
Taught Risks In Class	-0.235	Sports	-0.570
	(0.075)		(0.066)
Predicted (black)	0.1758		
Predicted (white)	0.2030		
# of Observations	8241		

Table 2: Results using detailed TAPS data.

Variable	MTF w/peers	TAPS w/peers	Variable	MTF w/peers	TAPS w/peers
Intercept	-1.853	-5.378	Male	-0.094	0.032
	(0.060)	(0.612)		(0.015)	(0.074)
Age	0.013	0.162	Parent HS	0.123	0.135
	(0.016)	(0.034)		(0.027)	(0.121)
Parent College	0.096	0.299	Church	-0.593	-0.594
	(0.015)	(0.087)		(0.017)	(0.075)
Urban	-0.292	-0.205	Suburban	-0.013	0.088
	(0.021)	(0.105)		(0.016)	(0.091)
Disp. Inc.	0.003	0.003	Midwest	0.101	0.182
	(1.39e-4)	(6.16e-4)		(0.021)	(0.106)
South	-0.106	0.047	West	-0.167	0.085
	(0.020)	(0.106)		(0.023)	(0.116)
% Friends Smoking	3.858	5.454	% Friends Smoking^2	-0.753	-1.268
	(0.111)	(0.411)		(0.101)	(0.387)
No Friends	1.589	0.017	Price	-0.177	
	(0.032)	(0.426)		(0.031)	
Predicted (black)	0.2547	0.1222			
Predicted (white)	0.3266	0.2129			
Social multiplier	2.5082	1.7655			
(std. error)	0.2141	0.1968			
# of Observations	120046	8241			

Table 3: Results w/peer behavior.

Variable	MTF w/peers	MTF no peers	Variable	MTF w/peers	MTF no peers
Intercept	-1.865	-0.356	Black	0.631	1.539
	(0.061)	(0.052)		(0.156)	(0.160)
Other Race	-0.348	-0.363	Male	-0.108	-0.181
	(0.022)	(0.020)		(0.015)	(0.013)
Age	0.031	0.075	Parent HS	0.050	-0.006
	(0.016)	(0.014)		(0.028)	(0.025)
Parent College	0.061	-0.007	Church	-0.582	-0.735
	(0.015)	(0.014)		(0.017)	(0.016)
Urban	-0.087	-0.045	Suburban	0.016	0.004
	(0.022)	(0.020)		(0.016)	(0.015)
Disp. Inc.	0.003	0.004	Midwest	0.090	0.012
	(1.41e-4)	(1.3e-4)		(0.021)	(0.019)
South	-0.022	-0.081	West	-0.169	-0.467
	(0.020)	(0.019)		(0.023)	(0.021)
% Friends Smoking	3.603		% Friends Smoking^2	-0.559	
	(0.112)			(0.101)	
No Friends	1.687		Price	-0.031	0.041
	(0.032)			(0.032)	(0.029)
Price * Black	-1.253	-2.060			
	(0.118)	(0.124)			
Predicted (black)	0.1514	0.1539			
Predicted (white)	0.3402	0.3405			
Social multiplier	2.3067				
(std. error)	0.1799				
# of Observations	120046	120046			

Table 4: Results with differential price response.

Variable		Variable	
Intercept	0.436	Male	-0.146
	(0.133)		(0.013)
Age	0.049	Parent HS	-0.429
	(0.014)		(0.131)
Parent College	-0.419	Church	-0.739
	(0.083)		(0.016)
Urban	0.523	Suburban	-0.033
	(0.108)		(0.014)
Disp. Inc.	0.002	Midwest	0.020
	(7.11e-4)		(0.019)
South	-0.202	West	-0.443
	(0.018)		(0.021)
Price	-0.666	Price * Disp. Inc.	0.001
	(0.096)		(5.06e-4)
Price * Parent HS	0.378	Price * Parent College	0.326
	(0.096)		(0.058)
Price * Urban	-0.585		
	(0.076)		
Predicted (black)	0.3090		
Predicted (white)	0.3190		
# of Observations	120046		

Table 5: Results with income-specific price response.

Variable	YRBS Baseline	Variable	YRBS Baseline
Intercept	-2.347	Male	0.194
	(0.142)		(0.017)
Age	0.080	Parent HS	0.093
	(0.007)		(0.031)
Parent College	0.054	No Parent Info	-0.082
	(0.023)		(0.031)
Hispanic	0.065	Price	-0.039
	(0.021)		(0.047)
Predicted (black)	0.2766		
Predicted (white)	0.2792		
# of Observations	65117		

Table 6: Results with YRBS data and state-specific prices.

C Proofs of propositions

C.1 Proposition 2.1

The right side of equation (6) is a continuous mapping from the unit interval to itself. By Brower's fixed point theorem, an equilibrium exists.

C.2 Proposition 2.2

Let $f(s) \equiv E_G(\Lambda(\beta X + \gamma(s)))$ First, note that f is continuously differentiable in s, and that for all G, $f(0) > 0$ and $f(1) < 1$.

First we prove by contradiction that if $f'(s) < 1$ at every equilibrium s, then there is only one equilibrium s. First, suppose that $f'(s) < 1$ for all equilibria s and that there is a continuum of equilibria (all $s \in [s_1, s_2]$ are equilibria). If so, then $f'(s) = 1$ for all $s \in (s_1, s_2)$, and we have a contradiction. Next, suppose that there is a discrete set of equilibria with multiple elements and that $f'(s) < 1$ for every equilibrium s. Select two equilibria s_1 and s_2 such that $s_1 < s_2$ and there is no equilibrium between them. Since $f'(s_1) < 1$ there exists (by the definition of a derivative) $\epsilon_1 > 0$ such that $f(s_1 + \epsilon_1) - (s_1 + \epsilon_1) < 0$. Since $f'(s_2) < 1$, there exists $\epsilon_2 > 0$ such that $f(s_2 - \epsilon_2) - (s_2 - \epsilon_2) > 0$. By the intermediate value theorem, there must exist some $s_3 \in (s_1 + \epsilon_1, s_2 + \epsilon_2)$ such that $f(s_3) = s_3$, and we have a contradiction. Therefore, if $f'(s) < 1$ for all s, equilibrium is unique.

Next, let $\sigma = \Lambda(\beta X + \gamma(s))$. Then $s = E_G(\sigma)$. We have:

$$\begin{aligned} f'(s) &= E_G\left(\gamma'(s)\,\Lambda'(\beta X + \gamma(s))\right) \\ &= \gamma'(s)\,E_G\left(\Lambda'\left(\Lambda^{-1}(\sigma)\right)\right) \end{aligned}$$

Now we can show through algebra that

$$\Lambda'\left(\Lambda^{-1}(x)\right) = x - x^2$$

Suppose we wished to select a distribution G of σ to maximize $f'(s)$ subject to the constraint that $E_G(\sigma) = s$. Since $\gamma'(s)$ is a constant, we can drop it. We wish to maximize

$$\int_0^1 \sigma - \sigma^2 dG(\sigma)$$

subject to the constraint that:

$$\int_0^1 \sigma dG(\sigma) = s$$

The solution to this maximization problem is that $\sigma = s$ with probability one, which implies that:

$$E_G\left(\gamma'(s)\,\Lambda'(\beta X + \gamma(s))\right) \leq \gamma'(s)\left(s - s^2\right)$$

This in turn implies that equilibrium is unique for every G if:

$$\gamma'(s)\left(s - s^2\right) < 1$$

for all s.

C.3 Corollary 2.3

In this case, equilibrium is unique if $\gamma s(1-s) < 1$. Since the quantity $s(1-s)$ is maximized (with value $1/4$) at $s = 1/2$, this condition can be changed to $\gamma * 1/4 < 1$ or $\gamma < 4$.

C.4 Corollary 2.4

In this case, equilibrium is unique if

$$(\gamma_1 + 2\gamma_2 s)s(1-s) < 1$$

This can be algebraically rearranged to get the result.

C.5 Proposition 2.5

First we have:

$$\begin{aligned}
\frac{d\bar{s}(G)}{dx} &= \frac{\partial f(G)}{\partial x} + \frac{\partial f(G)}{\partial s}\frac{d\bar{s}(G)}{dx} \\
&= \beta_x E_G\left(\Lambda'\left(\beta X + \gamma\left(s\right)\right)\right) + \\
&\quad \gamma'\left(s\right) E_G\left(\Lambda'\left(\beta X + \gamma\left(s\right)\right)\right)\frac{d\bar{s}(G)}{dx} \\
&= \frac{\beta_x E_G\left(\Lambda'\left(\beta X + \gamma\left(s\right)\right)\right)}{1 - \gamma'\left(s\right) E_G\left(\Lambda'\left(\beta X + \gamma\left(s\right)\right)\right)}
\end{aligned}$$

Dividing by:

$$\frac{\partial f(G)}{\partial x} = \beta_x E_G\left(\Lambda'\left(\beta X + \gamma\left(s\right)\right)\right)$$

we get:

$$m(G) = \frac{1}{1 - \gamma'\left(s\right) E_G\left(\Lambda'\left(\beta X + \gamma\left(s\right)\right)\right)}$$